

PD 09-APR-1998.
 XX
 PF 30-SEP-1997; 97WO-US17555.
 XX
 PR 30-SEP-1996; 96US-0026855.
 XX
 PA (EXSE-) EXSEED GENETICS LLC.
 XX
 PI Guan H, Keeling P;
 XX
 DR WPI; 1998-240100/21.
 XX
 DR P-PSDB; AAW56485.
 XX
 PT Hybrid polypeptide comprising starch-encapsulating region and
 PT protein - useful for, e.g. producing protein(s) resistant to
 PT degradation by stomach acids
 XX
 PS Example 2; Page 32-34; 156pp; English.
 XX
 CC The sequence is that of the waxy gene which codes for starch
 CC synthase. It can be used in the production of a hybrid
 CC polypeptide comprising a starch-encapsulating region (SER) fused
 CC to a payload protein. The hybrid polypeptide can be used to make
 CC modified starches comprising the payload protein, selected from,
 CC e.g. hormones, growth factors, antibodies, enzymes, dyes,
 CC immunoglobulins, etc. The modified starch can also be used
 CC to provide grain feeds enriched in amino acids. By encapsulating
 CC the payload protein in starch, it is more resistant to
 CC degradation by stomach acids.
 XX
 SQ Sequence 2542 BP; 610 A; 665 C; 693 G; 574 T; 0 other;
 Query Match 29.6%; Score 388.4; DB 19; Length 2542;
 Best Local Similarity 59.1%; Pred. No. 1.1e-56;
 Matches 795; Conservative 0; Mismatches 461; Indels 90; Gaps 4;

QY 5 TGGACATCGTGTGTTGCTGAGTCCGCGCTTGTTCAGAGCGCGCGCTGGCGG 64
 DB 700 TGAAGCTCGTGTCTCGCGCGCGAGATGGCCCTTGGAGCAAGACCGCGCTCGTG 759
 QY 65 ATGTGACTGTGGCTGCTATTGAGCTGTGCAAGCGCGCCACCGCTCATGACCATG 124
 DB 760 AGTCTCTGGTGGCTCCCTCCCTGCCATGGCTGCGNATGGCCACAGGTCATGTTGATCT 819
 QY 125 CCGCTGCTTACGACAGTACGTGAGCGCTGGGACACCTCGGTGTCGTGGACATCGG 184
 DB 820 CTCTCTCGGTACGACAGTACAGGACGCTTGGATACCAAGCGTTGGCTGAGATCAAGG 879
 QY 185 -----GCGAGAGTCCGCTACTTCCACTCCATCAAGAGGCGGTGCACCGG 232
 DB 880 TTGCAAGAGGTACGAGAGGTTGAGGTTTTCATTTGCTACAGCGTGGAGTCGACCGTG 939
 QY 233 TGTGGATTGACCAACCGCTGTTCTGCGGCAAGTCTGGGCAAGACCGCTCCAACTGT 292
 DB 940 TGTTCATGACCATCGTATTCCTTGGAGAAGTTTGGGGAAGACCGGTGAGAATCT 999
 QY 293 AGGCGCCCGCTCGCGCGCTGACTACTGTGGAACAACCAAGCGCTTCGCCCTGTTCTGCA 352
 DB 1000 ACGGACCTGACACTGAGTGTGATTACAAAGACAACAGATGGGTTTCAGCCCTCTTTGCC 1059
 QY 353 AGGCGCTATTGAGGTGCGCGCTGCTGCCCTTCGGCCCC----- 393
 DB 1060 AGGAGCACTGAGGCTCTAGATCTCTAAACCTCAACCAACCAACCACTACTCAAGAGAA 1119
 QY 394 -----GGCGAGACTCGCTTCTGCTGGCCAGCAAGTGGCACTCGCGCTGGCGCTGC 448
 DB 1120 CTTATGTTGAGGATGTTGTTGCTGCTGCAAGCACTGGCACACTGGCCCACTGGCAGCT 1179
 QY 449 TGTGAGGACGAGTACCAAGCGGCGAGTTCACCAAGGCGCAAGTGGTGGTGGTGA 508
 DB 1180 ACCTGAAGAACAACATACCAAGCGCAATGTCATACAGGAATGCAAGGTTGCTTTCTGCA 1239
 QY 509 TCCACACATCGCTTCCAGGGCGCGATGTGGGAGGAGGCTTTCAGGAGACACGAACTGC 568

DB 1240 TCACAAATCTCTCTACAGGGCGGTTTCGCTTTTCAGGATTACCTCGAGCTGAACCTCT 1299
 QY 569 CCAGGCGCGCTTTGACAAGCTGGCGCTTCTCGGACGCGCTATGCCAAGTTTACACTGAGG 628
 DB 1300 CCGAGAGTTTCAGTCTATCTTCGATTTCATCAGACGGGTATGAC----- 1343
 QY 629 CCACCCCATGAGGAGGAGGAGAGAGCCCGCTGACGGAAAGACTTACAAGAAGATCA 688
 DB 1344 -----ACGCGGTGGAGGCGAGGAATCA 1368
 QY 689 ATGGCTGAAGGTGGCATTATCGCCGCGACAAAGCTGGTGTGACTGTGCGCCCACTAG 748
 DB 1369 ACTGATGAAGCGCGAATCTTGAAGCGGACAGAGTGTCTACCGTGTAGCCCGTACTAG 1428
 QY 749 CGACCCAGATCGTCCGATGCGCGCGGTGTGTGAGCTGGACCGCTCATCCGCGCA 808
 DB 1429 CGAGGAGTCTATCTCCGGATCGCAGGGGATCGGAGTCTGACACATCATCGCGCTCA 1488
 QY 809 AGGCAATTGAGGSCATTGTGAAGCGCATGACATTGAGGAGTGAACCCCAAGACGACA 868
 DB 1489 CCGCATACACCGCATCGTCAACCGCATGACGCTGACGAGTGGGATCTTAGCAAGACA 1548
 QY 869 AGTTCGTGTCTGCGCCCTAGCACCAAGAGCTCTACGCCGCAAGGCGCGCCCAAGG 928
 DB 1549 AGTACATACCGCCAAAGTACGACGCAACCAACCGCAATCGAGGCGAAGGCGCTGAACAAG 1508
 QY 929 AGGCGCTGAGGCGGAGTGGCGCTGCTGTGACCCCAACCGCCCTGTTTCGCGCTTCA 988
 DB 1609 AGCGTTGAGGCGGAGCGGGTCTTCCGGTGCAGAGAAATCCCATGATCGCGTCA 1668
 QY 989 TCGGCGCTGAGGAGGAGAGGAGGTGTGACATCATCTTGGCGCCCTGCGCCCAAGATCC 1048
 DB 1669 TCGGAGGCTGAGGAGCAAGAGGCGCTGACGCTCATGGCGCGCCATCCCGGAGTCA 1728
 QY 1049 TGCCACCCCAAGTGTGACATCGCATCTGTGATACCGCAAGCGCGCTTACGAGAAGC 1108
 DB 1729 TG---CAGGAGGAGCTGCAGATCGTTCTTCTGGGTACTTGAAAGAGAGTTCGAGAAGC 1785
 QY 1109 TGTGAACGCTATCGGCACCAAGTACAAGGCGCGCCGCAAGGCGGTGGTCAAGTCTCGG 1168
 DB 1786 TGCTCAAGCATGAGGAGAGTATCCCGGCAAGGTGAGGCGGTGGTGAAGTCAACG 1845
 QY 1169 CGCCCTGGGCACTGCTACCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1228
 DB 1846 CGCGCTGTGCTATCTCATCATGTCGCGGAGCGGAGTGTGCTGCGCGTCCCGAGCGCTCG 1905
 QY 1229 AGCCTGCGGCTGATCCAGCTGACGCGCATGCTACGCTACGCTACGCTGCGCGTGTAGCCT 1288
 DB 1906 AGCCCTGTGACTCTCCAGCTGAGGAGTACGAGTACGAGTACGAGTACGAGTACGAGT 1965
 QY 1289 CCACCGCGCGCTGTGTCGACACCGTC 1314
 DB 1966 CCACCGGTGGCTCGTGGACACCGTC 1991

RESULT 4
 ABK53210
 ID ABK53210 standard; DNA; 2161 BP.
 XX
 AC ABK53210;
 XX
 XX 14-AUG-2002 (first entry)
 XX
 DE Potato granule-bound starch synthase I (GBSSI) DNA sequence.
 XX
 KW Starch; amylose; cereal crop; commodity starch; thickened foodstuff;
 KW bread; baking; pasta; potato; granule-bound starch synthase I; GBSSI; ds.
 XX
 OS Solanum tuberosum.
 XX
 PN WO200218606-A1.
 XX

XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132048.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.


```

PR 16-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      20.6%; Score 271.2; DB 21; Length 2115;
Best Local Similarity 54.3%; Pred. No. 5,le-37;
Matches 722; Conservative 0; Mismatches 518; Indels 90; Gaps 5;

QY 20 TTGCTGTGAGTCCGCCCTTGGTCCAAAGACGGCGGCTCGGGCGATGTGACTGGTGCC 79
DB 415 TTGGAGCTGAAGTTGGTCCATGGAGTAAACTGGTGGTCTTGGTGTATGTTCTCGGTGTC 474

QY 80 TGCCTATTGAGCTGTCAAGCGCGCCACCGGTCATGACCATTCGCCCTCGCTACGACC 139
DB 475 TACCTCCAGCTCTGTCTGCTAGAGCCACCGGTGTGATGACAAATTTGCTCGGTATGACC 534

QY 140 AGTAGCTGAGCCTGGGACACCTCGGTGGTCTGGTCAAGTTCAGATCAAGTTGGGGATAAAGTTG 187
DB 535 AATATAAGATGCTTGGGACACTTGTGTGTGGTTCAGATCAAGTTGGGGATAAAGTTG 594

QY 188 AGAAGTCGGCTACTTCCACTCCATCAAGAAGGGCGTCCACGGGTGTGGATGTACCAACC 247
DB 595 AGAATGTTCGTTTCTTCCATTGCTACAAACGAGGAGTTGATCGTGTCTTTTGTGACCATC 654

QY 248 CCTGGTTCCTGGCAAGTCTGGGCAAGCGGCTCCAAAGCTGACGGCCCGCCCTCGG 307
DB 655 CAATCTTCTTGTCTAAGTTGTGGCAAAACAGGATCCAAATCTATGCTCTATAACTG 714

QY 308 GCGTGTACTACTGGACACCAACCAAGCTTCCGCCCTGTTCTGCAAGCGCGCTATTGAGG 367
DB 715 GAGTAGACTACAATGACAAACCACTCCGGTTCAGTTGTTGTGTGTCAGGCTGCTCTTGAGG 774

QY 368 CTGCCCGGCTGCTGCCCTTCGGCCCGCGGAGGACTGC----- 405
DB 775 CACCACAGGTTCTGAACCTGAACAGCAGCAAGTACTTCTCTGGACCAATATGTTGAAGATG 834

```

```

QY 406 --GCTCTCTGTGGCCAAAGCACTGGCACTCCGCCCTGTGTGGCCGCTCCTGCTGAAGCAACGAGT 463
DB 835 TAGTCTTTTGTGGCAATGACTGGCACACTGCTCTACTTCCATGTACTTCAATCTAATGT 894

QY 464 ACCAGCCCAAGGGCCAGTTCCACCAAGCCCAAGTCGGTGTGGCTATCCCAACATCCGCT 523
DB 895 ATCAATCCCGCGGAGTCTACATGAATGCAAAAGGTGCTTCTGCTATTCAACAATACGCT 954

QY 524 TCCAGGCGCCATGTGGGAGGAGGCTTCAAGACACCAAGAGCTGCCCGAGCGCCCTTG 583
DB 955 ACCAGGGAAGATTT-----GCCCTTG 975

QY 584 ACAAGCTGGCCTTCTCGGACGGCTATGCCAAGGTTTACACTGAGGCCACCCCATGAGG 643
DB 976 ATGACTATTCCCTTCTCAACTTGCCTCATGCTTTAAAGTCTTTCGACTTCATGGAGC 1035

QY 644 AGGACGAGAAGCCCGCTGACGGGAAAGACCTTACAAAGAGATCAACTGCTGAAAGGTTG 703
DB 1036 GGTATGAAAAGCC-----AGTAAAGGACGGAATTAACCTGGATGAAGGCTG 1083

QY 704 GCATTATCCGCGCCGACAAAGCTGGTCACTGTGTCGCCCACTACGCGACCGGATCGCTG 763
DB 1084 CAATTCTGGAAGCTACCGGTGCTTAAACAGTTAGTCCACTACTACGCTCAAGAACTCATCT 1143

QY 764 CCGATGCCCGCGGCTGTGGAGCTGGACACACCGCTCATCCGCGCAAGGGCATTGAGGGCA 823
DB 1144 CTGGAGTTGATAGAGCGGTGGAATTGCATAAATACTTCAATGAAACAGTTTCGGGA 1203

QY 824 TTGTGAACGGATGGACATTTAGGAGTGGAAACCCCAAGACCCGACAAAGTTCCTGTCTCCG 883
DB 1204 TTATTAATGAATGGATGTTCAAGAAATGGAACCCGCTACTGACAAAGTACATGATATCA 1263

QY 884 CCTACGACCAAGACCGGCTTACGCCGCGAAGCGCCGCCAGGAGGCGCTTCAGGCGG 943
DB 1264 AATACGATATTACCCTGTTACAGATGCTTAACCATTTGATCAAGAAGACACTTCAGGCTG 1323

QY 944 AGCTGGGCTTGTGTGGACCCACCGCCCTGTTCTGCTTCATCGGCGCGCTGGAGG 1003
DB 1324 CTGTTGGACTTCCGTTGGACAGGGATGTCGCGTATCGTTCATAGGAGATTGGAGG 1383

QY 1004 ASCAGAAAGGTGTGACATCATCTTGGCGCGCCCTGCGCCCAAGATCTTGGCCACCCCAAG 1063
DB 1384 ASCAGAAAGGTTCATGATATTTAGTGAAGACTATTTCCAAAGTTCATG---GSCCTCAATG 1440

QY 1064 TGCACATCCCATCTGGGTACCGCAAGCGCCCTTACGAGAAGCTGGTGAACGCATCG 1123
DB 1441 TTCAGATGTTATCCTTGGGACTGGAAAGAAAGATGAGGCTCAGATCTTTGAACATG 1500

QY 1124 GCACCAAGTACAAAGGCGCGCCCAAGGGGCTGGTCAAGTTCTCGGCGCCCTGGCGCACA 1183
DB 1501 AAGAGAAAGTCCCAAGGAGGCGGTGGAGTGGCGAAATTCACGTTGCCATTGGCTCATG 1560

QY 1184 TGCCTACCGCGCGCGGACCTTCATGCTGGTCCCTCGGCTTCGAGCCCTCGGCGCTGA 1243
DB 1561 TGATCACTGCTGGAGCTGACITTCATGTTGCCCAAGCAGGTTTGGCGGTGTGCTCTCA 1620

QY 1244 TCCAGCTCACGCCCATGCACTACGCTACGCTGCGCCGCTGTGCTCCCTCCACCGCGGCTGG 1303
DB 1621 TTCAGCTGCACGCAATGAGATGATGAAACCGTCCCTTATGTTGTCATCTACTGTGGACTTG 1680

QY 1304 TGGACACCGT 1313
DB 1681 TGGACACTGT 1690

```

```

RESULT 6
ABK53215
ID ABK53215 standard; cdna; 2183 BP.
XX AC ABK53215;
XX DT 14-AUG-2002 (first entry)

```

xx Canna edulis granule-bound starch synthase I (GBSSI) cDNA sequence.
 DE Starch; amylose; cereal crop; commodity starch; thickened foodstuff;
 KW bread; baking; pasta; potato; granule-bound starch synthase I; GBSSI; ss.
 XX Canna edulis.
 OS WO200218606-A1.
 PN 07-MAR-2002.
 XX 28-AUG-2000; 2000WO-US23494.
 XX 28-AUG-2000; 2000WO-US23494.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Broglie K, Lightner J;
 XX WPI; 2002-404556/43.
 XX Novel starch isolated from cereal crop with a modification of amylose
 PT content and/or amylose weight distribution relative to commodity starch
 PT from same species of the cereal crop, for preparing thickened
 PT foodstuffs
 XX Example 1; Page 37-38; 44pp; English.
 XX The present invention relates to a new starch comprising amylose, 50%
 CC where the starch is isolated from a cereal crop, and comprises a
 CC increase in average molecular weight relative to commodity starch from
 CC the same species of the cereal crop from which starch was isolated.
 CC The starch of the invention is useful for preparing thickened foodstuffs
 CC by combining the starch, a foodstuff and water and cooking the resulting
 CC composition as necessary to produce a thickened foodstuff. The invention
 CC is also useful for preparing foodstuffs such as breads, baked goods,
 CC pastas, etc. The present nucleic acid sequence represents the Canna
 CC edulis granule-bound starch synthase I (GBSSI) cDNA sequence in clone
 CC ectic.pk007.015 that was used in the methods of the invention for
 CC generation of a full-length sense Canna edulis GBSSI construct.
 XX Sequence 2183 BP; 636 A; 428 C; 516 G; 603 T; 0 other;
 SQ
 Query Match 16.5%; Score 217.2; DB 24; Length 2183;
 Best Local Similarity 51.1%; Pred. No. 5.9e-28;
 Matches 688; Conservative 0; Mismatches 568; Indels 90; Gaps 4;
 QY 5 TGCACATCGTATGTTGCTGTGAGTGGCGCCCTTGGTCCAAAGAGGGCGCGCTGGGGC 64
 DB 357 TGAACCTGGGGTTTGTGTTGTGAGTGGTCCATGAGCAAAACTGGGGGCGCTTGGCG 416
 QY 65 ATGTGACTGGTGGCTGCTATTGAGTGGTCAAGCGGCGCCACCGGTCATGACCATG 124
 DB 417 ATGTTCTTATAGATTGCCACCTGCTATGCGTGAATTTGGGCACAGGGTCATGACCGTGG 476
 QY 125 CCCTCGCTACGACAGTACGCTGCGCTGGGACACCTCGGTGGTTCGTGGACAT----- 179
 DB 477 CGCCAGATGACCAATATAAGATATCTGGATACAGTGTCCCGATTGAGTTAAAG 536
 QY 180 -----CATGGCGAGAAGTCCGCTACTTCCACTCCATCAGAGAGGGCGTGCACCGG 232
 DB 537 TTGGGATTAAGATTAAACTGTCCGCTTCCACTGCTACAAAGGGGAGTTGATCGG 596
 QY 233 TGTGGATTCACACCGCTGCTTCCGCAAGTCTGGGCGAAGCGCGCTCCAACTGT 292
 DB 597 TTTTGTGATCACCTATGTTTCTCGAGAAGTTTGGGGAAACAGAGGAAATAT 656
 QY 293 AGGCCCCCGCTCCGCGCTGACTACCTGGCAACCAAGCGCTTCGCCCTGTTCTGCA 352
 DB 657 ATGTCCTGTTACAGGAACAGATTATGACAGCAATCAACTAAGATTCAGCCTTTTGTGC 716
 QY 353 AGGCCCTATTGAGGCTGCCCCGCTGCTGCCCTTCGGCCCC----- 393

DB 717 TGGCAGCTCTGGAAGCTCCAAAGACTTCTAAATCTCAACAACAGCAAAATCTATTCTGGAC 776
 QY 394 -----GGCGAGGACTGCTTCTCGTGGCGCAAGCTGGCACTCGCGCTGGTGGCGCTCC 448
 DB 777 CATATGGAGATGATGTTGTGTTTATGCAACAGATGGCATTTCTGCTACTGCGCTGCT 836
 QY 449 TGTCTGAAGGACGAGTACCAAGCCCAAGGGCCAGTTCCACCAGGCCAAGTGGTGGTGA 508
 DB 837 ACTTCAAAACTATGTACCAATCACATGTTATTACATGAATGCTAAGGTTGATTTTGA 896
 QY 509 TCCACAACATCCCTTCCAGGGCGCATGTGGAGGAGGCTTCAAGGACACGAAGCTGC 568
 DB 897 TTCATAATATTGCTTACCAGGGCGCATTTGCTTTTCGAGCTTTGAACCTCTTAATCTCC 956
 QY 569 CCAGGCGCGCTTGTACAAGCTGGCTTCTCGGAGGGCTATGCCAAGGTTTACACTGAGG 628
 DB 957 CCAATAAATTTAAATCTTCAATTGATTCATGATGATATGACAACT----- 1006
 QY 629 CCACCCCATGAGGAGGACGAGAGAGCCCGCTGACGGGAAGAGCTACAAGAAGATCA 688
 DB 1007 -----GTGAAGGAAGGAATAA 1025
 QY 689 ACTGCTGAGGGTGGCATTTATCGCGCGCAAGCTGGTGTGCTGCTGCCCAACTAGC 748
 DB 1026 ATTGGATGAAGCTGGAATTAATAGAAATGATAGGTGCTTTGACCGTGAGCCCATATATG 1085
 QY 749 CGACCGAGATCGTCCGATGCGCGCGGTGTGTCAGCTGGACACCGCTCATCCCGCCA 808
 DB 1086 CCCAAGAGCTTGTCTCAGGGGTAGAGAAGGTCTTCAGTTGGCAATATCTTCGCGCATGA 1145
 QY 809 AGGGCATGAGGCGCATTTGAACGGCATGACATTCAGGAGTGGAGACCCCAAGACCGACA 868
 DB 1146 AAACCATCTGTGGAATAGTAAATGGGATGGACACACCGAGGAGTAACTCAATTAACAGA 1205
 QY 869 AGTTCCTGCTGCGCCCTAGCAGACCAAGAGGTCTACCGSCAAGGCGCGCGCAAGG 928
 DB 1206 AATATATTTCTACAACACTAGATGCAACAACTGATGATGATGCAAAACCTCTCTGTAAG 1265
 QY 929 AGCGCCCTGAGGCGGAGCTGGGCTGCTGTTGGACCCCAAGCGCCCTGTTTCGCTTCA 988
 DB 1266 AAGCTTTGCAAGCTGAGTGGGCTGCTTAAACAAACAAAGCTTGTTCGCGCTTG 1325
 QY 989 TGGCGCGCTGGAGGACAGAGGGTGTGGACATATCTTGGCGCGCTGCGCCCAAGATCC 1048
 DB 1326 TTGGAAGACTAGATGACAGAAAGGCTCAGACATTTAGCTGACGCAATTCAGAGACTTC 1385
 QY 1049 TGGCCACCCCAAGGTGCAGATCGCCATCTGCGGTACCGSCAAGGCGCGCTACGAGAAGC 1108
 DB 1386 TTTGTGAG---AATGTTCAAGTGAATAGTACTTGGGCACTGSCAAGAGAGTGGAGAGTG 1442
 QY 1109 TGTGAACGCCATCGGCACCAAGTACAGGGCGCGCGCCCAAGGCGGTGGTCAAGTTCTCGG 1168
 DB 1443 AACTTACATTAATTGAGGAAATGTTTCCAGACAAATTCAGAGCACATCTCAAAATCAACG 1502
 QY 1169 CGCCCGCTGGGCACATGCTCACCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1228
 DB 1503 TTTCTTTAGCTCATGCAATCATGCGAGGAGCTGATATCTTGTATTTCCAAGCAGATTCG 1562
 QY 1229 AGCCCTGCGSCCTGATCCAGCTGCACGCCATGCACTACCGTACCGTGGCGGTGAGCCT 1288
 DB 1563 AACCTTGTGGCTTCATTCAGCTTCAGGCCCATGGATATGGAATCTCCCTATGTTAGCA 1622
 QY 1289 CCACCGCGCGCTTGGTGCACACCGTC 1314
 DB 1623 CCACCTGTTGACTGTTGACACTGTC 1648

RESULT 7
 AAC86435
 ID AAC86435 standard; cdna; 2807 bp.
 XX
 AC AAC86435;

US5824790-A.

9

US5824790-A.

Db 1214 AYGAITYTYGNAAYTYTGAYTYTCCNGARCAATAYATHGAYCAVITYAARYTNIAYGAYA 1273
 QY 605 GCTATGCCAAGGTTTACACTAGGCGCACCCCATGGAGGAGGAGGAGAGCCGCCCGCTGA 664
 Db 1274 AYATHGGNGGAYCAYWSNAYTYTNTYGC----- 1304
 QY 665 CGGGAAGACCTTAAAGAAGATCAACTGGCTGAAGGTGTGCAITATCGCGCGGCAAGC 724
 Db 1305 -----NGCNGNYTNAARACMGNGAYMGNG 1330
 QY 725 TGGTGACTGTGCGCCCACTAGCGACCGAGATCGCTGCCGATGCCGCGCGGCTGTGG 784
 Db 1331 TGTNACNGTNNWNAAYGNTATATGCGARYTNAARACNWSNGARGNGGNTGGGNY 1390
 QY 785 AGCTGACACCGCTATCCGCGCAAGGCA---TTGAGGCGATTGTGAACGGCGATGACA 841
 Db 1391 TNCAYGAYATHAAYCARAYGATGGAARYTNCARGNATHGTNAAAYGGNATHGAYA 1450
 QY 842 TTGAGAGTGGACCCCAAGACCGCAAGTTCCTGTCTGCGCCCTACGACCAAGC--- 897
 Db 1451 TGWSNGARTGGAAYCNGCNGTNGAYGTNCAVYTNCAVWSNGAYGAYTAYACNAAYTAYA 1510
 QY 898 -----AGCCTCTACGCGCGGAGGCGCGCCGCAAGGAGGCGCTTGCAGCGCGAGCTGG 949
 Db 1511 CMNTYGARACNTYNGAYACNGGNAARMGNCARTGYAARGCNGNTNCAARMGNCARYTNG 1570
 QY 950 GCGTGCCTGTGGACCCCGCGCCCGCTGTGCGCTTCTATCGCGCCCTGGAGGAGCAGA 1009
 Db 1571 GNYTNCARCTNMGNGAYGAYGTNCCNYNATHGGNTTYATHGGNMGNTNGAYCAYCARA 1630
 QY 1010 AGGTGTGGACATATCTTGGCGCGCTGCCCAAGATCTGCGCACCCCGCAAGGTGSCACA 1069
 Db 1631 ARGNGTNGAYATHAGCNGAYGNCNATHCAVYTGGAH---GCNGGNCARGAYGTNCARY 1687
 QY 1070 TCGCATCTGGTACCGGACGCGCGCTTACGAGAAGTGTGTGAACGCCATCGGCACCA 1129
 Db 1688 TNGTNTATYTNNGNACNGNGGNGGNGAYTNGARGAYATGTYNMGNTTYGARWSNG 1747
 QY 1130 AGTAAAGGCGCGCCCAAGGCGGTGTCAGTTCCTGCGCGCCCTGCGCGCTGATCCAGC 1189
 Db 1748 ARCAVWSNGAYARCTNMGNGCTGGTNGGNTTWSNGTNCNTYNGCNCAYMGNATHA 1807
 QY 1190 CGCGCGCGCGCTATGCTGTGCTGCGCTTGCAGCGCTGCGCGCTGATCCAGC 1249
 Db 1808 CNGCNGGNGCNGAYATHYTNATGCGNWSNMGNTTYGARCCNTYGYGNTYNAAYCARY 1867
 QY 1250 TGCACGCCATCACTACGTACCGTGCCTGCTGCTAGCCTCCACGCGCGCGCTGTGCGACA 1309
 Db 1868 TNYAGCNATGGCNTAYGNGACNGTNCNGTNGTNCAYGCGNTGNGGNGNTNMNGNAYA 1927
 QY 1310 CCGT 1313
 Db 1928 CNGT 1931

RESULT 12

ID AAC86412 standard; cdna; 2107 BP.
 AC AAC86412;
 AC AAC86412;

DT 01-MAR-2001 (first entry)

DE Wheat starch synthase II coding sequence SEQ ID NO: 5.

XX Wheat; starch synthase; SSII; SSIII; starch content; starch synthesis;
 KW food product; adhesive; ss.
 XX Triticum aestivum.

OS WO200066745-A1.
 PN 09-NOV-2000.

XX

PF 28-APR-2000; 2000WO-AU00385.
 PR 29-APR-1999; 99AU-0000052.
 XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
 PA (GOOD-) GOODMAN FIELDER LTD.
 PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
 XX Morell M, Li Z, Rahman S, Appels R;
 PI WPI: 2000-647602/62.
 XX P-PSDB; AAB37568.
 DR Nucleic acid molecules encoding wheat starch synthase (WST)-I and
 DR WST-II, useful in modifying plant starch content and/or composition -
 XX Claim 1: Page 161-163; 21lpp: English.
 XX The present invention relates to novel protein and coding sequences from
 CC wheat. The proteins are wheat starch synthases, designated SSII and
 CC SSIII. These can be used in the modification of plant starch content or
 CC composition, and to screen plants to identify mutations which affect
 CC starch content and composition. The starch can then be used in food
 CC products, such as flour, and in films, coatings, adhesives, building
 CC materials and packaging materials.
 XX Sequence 2107 BP; 452 A; 566 C; 664 G; 425 T; 0 other;
 SQ

Query Match

Best Local Similarity 15.1%; Score 198.4; DB 21; Length 2107;
 Matches 388; Conservative 0; Mismatches 236; Indels 18; Gaps 3;

QY 687 CAACUGGCTGAAGGCTGCGATTATCCCGCCGACACAGCTGGTGTGCGCCCAACTA 746
 Db 933 CAACACTCTTCGCGCGCGCTGAAGATGGCGGACAGCTGTGCTGTGAGCGCGGTA 992
 QY 747 CGCGACCGAGATCGCTGCGGATGCCCGCGGCTGGAGCTGGACACCGTCATCC---G 803
 Db 993 CCGTGTGGAGCTGAAGACGCTGGAGCGCGCTTACGACATCATAGGCGAGA 1052
 QY 804 CGCCAAAGGCTATGAGGCGATTGTGAACGGCATGGACATTTGAGGAGTGAACCCCAAGAC 863
 Db 1053 CGACTGGAAGACCCCGCGCATGTCAACGCGCATCGACAACATGGAGTGAACCCCGAGT 1112
 QY 864 CGACAAGTTCCTGTCTGCGCCCTACGACCGGCTACCAACTTCTCCCTGAGGACGCTGGACTCCGG 911
 Db 1113 GGACGCCCACTCAAGTCGGACGGGTACCAACTTCTCCCTGAGGACGCTGGACTCCGG 1172
 QY 912 CAAGCGCGCGCCCAAGGAGGCGCTGACGCGCGAGCTGGGCGCTGCTGTGGACCCGACCGC 971
 Db 1173 CAAGCGCGAGTGAAGGAGGCGCTGACGCGCGAGCTGGGCGCTGAGGTCGCGCGAGCT 1232
 QY 972 CCCCCTGTTCGCTTCATCGCGCGCTGGAGGACGAGAGGCTGGACATCATCTGCGC 1031
 Db 1233 GCGCTGTCTGCTTCATCGCGCGCTGGAGGACGAGAGGCTGGAGATCATCGCGA 1292
 QY 1032 CGCCCTGCCCAAGATCTGCGCACCCCAAGTGCAGATCGCCATCTCTGGGTACCGGCAA 1091
 Db 1293 CGCCATGCCCTGGATCGTG---AGCCAGGAGTGCAGCTGGTGTGCTGGACACCGGCGC 1349
 QY 1092 GCGCGCTTACGAGAGCTGGTGAAGCGCATCGCACCAAGTACAAGGCGCGCGCAAGGG 1151
 Db 1350 CCAGGACCTGGAGAGCATGCTGCAGACTTCGAGCGGAGGACCCACACAGGTGCGCGG 1409
 QY 1152 CGTGTCAAGTCTTCGCGCGCGCTGGGCGCATGCTCACCGCGCGCGCGCTTCACTGCT 1211
 Db 1410 GTGGGTGGGTTCTCCGCTGCGCTGGCGCATCGCGGATCACGCGCGCGCGCGCTCT 1469
 QY 1212 GGTGCCCTCGCGCTTCGAGCGCTGCTCCAGCTGATCCAGCTGCACGCCATGCTAGGTAC 1271
 Db 1470 CATGCCCTCCCGGTTCTGCGCTGGGCTGAACCACTCTACGCCATGCGCTACGGCAC 1529

Query Match	15.0%;	Score 196.8;	DB 21;	Length 9024;
Best Local Similarity	60.3%;	Pred. No. 1.5e-24;		
Matches 387;	Conservative	0;	Mismatches 237;	Indels 18; Gaps
QY	687	CAACTGGCTGAAGGGTGGCAATATCCGCCGCACAGCTGGTACATGTGTGCCCACTA	746	
Db	7392	CAACTACTTCGCCCGCGGCTGAAGATGGCGACAGGTGTGCTGGTAGCCCGGGTA	7451	
QY	747	CGCGAOCGAGATCGCTCCCGATGCCGCCGCCGTGTGGAGCTGGACACCGTCATCC	803	
Db	7452	CCTGTGGAGCTGAAGACGGTGGAGCGCGCTGGGGCTTCACGACATCATACGCGCAA	7511	
QY	804	CGCCAAAGGCAATTGAGGGCAATTGTAAACCGCATGACATTGAGGAGTGGACCCAGAC	863	
Db	7512	CGACTGGAAGACCCGGCATCGTCAACGGCATGCACAACTATGGAGTGGAAACCCGAGT	7571	
QY	864	CGACAAGTTCTCTCTCGGCCCTACGACCAGAACAGCGTCT	911	
Db	7572	GGACGCCACCTCAAGTCGGACGGCTACACCAACTTCCTCGTAGGACGCTGGACTCCG	7631	
QY	912	CAGGCGCGCCCAAGAGAGCCCTGCAGCCACAGCTGGCCCTGCCTGTGGACCCCAACGC	971	
Db	7632	CAGCGCAGTGCAGAGGAGGCCCTGCAGCGCAGCTGGGCCCTGCAGTTCGCGCCGACGT	7691	
QY	972	CCCCCTGTTCCGCTTCATCGGCGCCTGGAGGACAGAAAGGTGTGGACATCATCTCTGC	1031	
Db	7692	GCCTGTCTCGGCTTCATCGGCGCCTGCAGCGGCAGAGGGCTGGAGATCATCGCGA	7751	
QY	1032	CGCCCTGCCCCAAGATCTTGGCCACCCCAAGGTGCAGATCGGCATCTCTGGGTACCGGAA	1091	
Db	7752	CGCATGCCCCGTGATCGTG---AGCCAGAGACTGCAGCTGGTGTGCTGGCAACCGGCG	7808	
QY	1092	GGCGGCTACGAGAAGCTGGTGAACGCATCGGCACCAAGTACAGAGGCCGCGCAAGGG	1151	
Db	7809	CCAGCACTGGAGAGCATCTCGGCACTTCAGCGGGAGCACCAACGACAGGTGCGCGG	7868	
QY	1152	CGTGTCAAGTTCTCGGCGCCCTTGGCGCAATGCTACCCGCCCGCGCCGATTCATGCT	1211	
Db	7869	GTGGGTGGGGTTCTCGCTGCGCTGCGCACCGATCACGGCGGGGCGGACGCGCTCT	7928	
QY	1212	GGTGCCTTCGCGCTTCAGCCCTCGGCTGATCCAGCTGCACGCCATGCACTACGCTAC	1271	
Db	7929	CATGCCCTCCCGTTCGACCGTGGGGCTGAACCACTCTACGCCATGGCTACCGCAC	7988	
QY	1272	CGTGGCGGTGGTAGCCTCCACCGCGGGCCTGGTGACACCGT	1313	
Db	7989	CGTCCCGCTGTGCAGCGGCTTGGGGCTCAGGCAACCGT	8030	

RESULT 15

AAC86410
 ID AAC86410 standard; cDNA; 2939 BP.

AAC86410;

XX
DT 01-MAR-2001 (first entry)

DE Wheat starch synthase II coding sequence SEQ ID NO: 1.

Wheat; starch synthase; SSII; SSIII; starch content; starch synthesis;
KW food product; adhesive; ss.

OS *Triticum aestivum*.

XX
PN
WO200066745-A1.

XX
PD 09-NOV-2000.

28-APR-2000; 2000WO-AU00385.

XX
PR 29-APR-1999; 99AU-0000052.

XX
PA (CSIR) COMMONWEALTH SCI & IND RES ORG.

PA (GOOD-) GOODMAN FIELDER LTD.
FA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
XX
PI
xx Morell M, Li Z, Rahman S, Appels R;
xx WPI; 2000-647602/62.
DR P-PSDB; AAB37566.
DR

PT Nucleic acid molecules encoding wheat starch synthase (WST)-I and
PT WST-II, useful in modifying plant starch content and/or composition -
XX
PS Claim 1; Page 149-152; 21pp; English.

CC The present invention relates to novel protein and coding sequences from
CC wheat. The proteins are wheat starch synthases, designated SSII and
CC SSIII. These can be used in the modification of plant starch content or
CC composition, and to screen plants to identify mutations which affect
CC starch content and composition. The starch can then be used in food
CC products, such as flour, and in films, coatings, adhesives, building
CC materials and packaging materials.

SQ Sequence 2939 BP; 580 A; 892 C; 946 G; 521 T; 0 other;

```
Query Match      14.7%; Score 193.6; DB 21; Length 2939;
Best Local Similarity 60.0%; Pred. NO. 5.3e-24;
Matches 385; Conservative 0; Mismatches 239; Indels 18; Gaps 3;
```

Search completed: June 3, 2003, 10:08:50

Job time : 232.966 secs

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2003, 09:33:29 ; Search time 1282.02 Seconds
(without alignments)
16599.542 Million cell updates/sec

Title: US-09-980-771A-6
Perfect score: 1314
Sequence: 1 gcctgacatcgtgatggt.....gcggcctggtcacacgctc 1314

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_esti:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_roo:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	469.8	35.8	473	10	AV640741
2	457.8	34.8	461	10	AV629488
3	444.4	33.8	458	10	AV622787
4	400.8	30.5	436	10	AV629117
5	395.6	30.1	511	10	AV394077
6	384.6	29.3	2147	11	AY109531

7	365.2	27.8	475	10	AW757933
8	337.8	25.7	418	10	AV644765
9	295.8	22.5	506	10	AV643281
10	278.4	21.2	702	14	BQ246353
11	272.8	20.8	560	10	AV628313
12	251.8	19.2	552	10	BE024926
13	250.2	19.0	517	10	AV641583
14	245.6	18.7	726	14	BQ804991
15	244.2	18.6	532	10	AV641989
16	235.8	17.9	513	10	AV631004
17	235.8	17.9	552	10	AV640546
18	234.8	17.9	524	10	AV641724
19	234.2	17.8	500	10	AV642869
20	231.8	17.6	510	10	AV395307
21	223.8	17.0	540	10	AV642722
22	211.8	16.1	486	10	AV642834
23	211.8	16.1	504	10	AV629198
24	210.8	16.0	502	10	AV644517
25	208.8	15.9	495	10	AV642565
26	207.8	15.8	696	12	BG351920
27	207	15.8	641	14	BQ247154
28	206.8	15.7	522	10	BE423625
29	206.4	15.7	729	12	BG599615
30	201.2	15.3	585	14	BQ246366
31	200.8	15.3	493	10	AV631838
32	199.8	15.2	492	10	AV629891
33	198.8	15.1	484	10	AV629266
34	198.8	15.1	489	10	AV629190
35	196.8	15.0	773	12	BG351175
36	196.2	14.9	734	12	BG350240
37	193.8	14.7	487	10	AV643763
38	193.6	14.7	385	10	AV393664
39	193.6	14.7	436	14	BQ245102
40	189.8	14.4	541	10	BE424431
41	189.8	14.4	631	13	BE531619
42	188.8	14.4	490	10	AV628140
43	187.8	14.3	476	10	AV644278
44	185.8	14.1	478	10	AW758015
45	185.8	14.1	502	10	AV631098

ALIGNMENTS

RESULT 1	AV640741	AV640741	473 bp	mRNA	linear	EST 15-DEC-2000
LOCUS	AV640741	AV640741	Chlamydomonas reinhardtii	5% CO2	Chlamydomonas reinhardtii	
DEFINITION	AV640741	AV640741	CDNA clone HCL021f06_r 5', mRNA sequence.			
ACCESSION	AV640741	AV640741.1	GI:10784069			
VERSION	AV640741	EST				
KEYWORDS	EST					
SOURCE	Chlamydomonas reinhardtii					
ORGANISM	Chlamydomonas reinhardtii					
REFERENCE	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas.					
AUTHORS	1 (bases 1 to 473)					
TITLE	Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K., Nakamura, Y. and Tabata, S.					
JOURNAL	Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii					
MEDLINE	DNA Res. 7 (5), 305-307 (2000)					
COMMENT	20539644					
CONTACT	Contact: Erika Asamizu					
	The First Laboratory for Plant Gene Research					
	Kazusa DNA Research Institute					
	Yana 1532-3, Kisarazu, Chiba 292-0812, Japan					
	Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.					
FEATURES	Location/Qualifiers					
Source	1..473					
	/organism="Chlamydomonas reinhardtii"					
	/strain="C9"					
	/db_xref="taxon:3055"					

```
/clone="HCL021f06_r"  
/clone_lib="Chlamydomonas reinhardtii 5% CO2"  
/note="Vector: pBluescriptII SK-; Site.1: EcoRI; Site.2:  
XhoI; The cDNA library was constructed from cells cultured  
in a medium with bubbling air containing 5% carbon  
dioxide"
```

```
BASE COUNT      76 a   162 c   146 g   89 t  
ORIGIN  
  
Query Match      35.8%; Score 469.8; DB 10; Length 473;  
Best Local Similarity 99.6%; Pred. No. 3.8e-81; Indels 0; Gaps 0;  
Matches 471; Conservative 0; Mismatches 2;  
  
QY 17 TGGTTGCTGCTGAGTGGCCCTTGTCCTCAAGACGGCGGCTGATGACTGGTG 76  
Db 1 TGGTTGCTGCTGAGTGGCCCTTGTCCTCAAGACGGCGGCTGATGACTGGTG 60  
  
QY 77 GCCTGCTATTGAGTGGTCAAGCGGGCCACCGCGTCATGACCATTCGCCCTGCTACG 136  
Db 61 GCCTGCTATTGAGTGGTCAAGCGGGCCACCGCGTCATGACCATTCGCCCTGCTACG 120  
  
QY 137 ACCAGTACGCTGACGCTGGGACACCTCGGTGTCGTGGACATCATGGCGAGAGGTCC 196  
Db 121 ACCAGTACGCTGACGCTGGGACACCTCGGTGTCGTGGACATCATGGCGAGAGGTCC 180  
  
QY 197 GCTACTTCCACTCCATCAAGAAGGGCGTGCACCGCGTGTGGATTGACCAACCCCTGGTTC 256  
Db 181 GCTACTTCCACTCCATCAAGAAGGGCGTGCACCGCGTGTGGATTGACCAACCCCTGGTTC 240  
  
QY 257 TGGCCAGGCTTGGGCAAGACCGGCTCCAAAGCTGTACGCGCCCGCTCGGCGCTACT 316  
Db 241 TGGCCAGGCTTGGGCAAGACCGGCTCCAAAGCTGTACGCGCCCGCTCGGCGCTACT 300  
  
QY 317 ACCTGGCAACCAACGAGCGTGCCTCTGTCGAAGGCGGCTATTGAGGCTGCCGCG 376  
Db 301 ACCTGGCAACCAACGAGCGTGCCTCTGTCGAAGGCGGCTATTGAGGCTGCCGCG 360  
  
QY 377 TGCTGCCCTTCGGCCCGCGGAGACTGCGTCTTCGTGGCCCAAGCATGGCACTCCGCCC 436  
Db 361 TGCTGCCCTTCGGCCCGCGGAGACTGCGTCTTCGTGGCCCAAGCATGGCACTCCGCCC 420  
  
QY 437 TGGTGGCCGCTCTGCTCAAGACGAGTACCAAGCCCAAGGCGCAGTTCACCAAG 489  
Db 421 TGGTGGCCGCTCTGCTCAAGACGAGTACCAAGCCCAAGGCGCAGTTCACCAAG 473
```

```
RESULT 2  
AV629488  
LOCUS      AV629488 Chlamydomonas reinhardtii 5% to 0.04% CO2 Chlamydomonas  
DEFINITION reinhardtii cDNA clone LCL059f09_r 5', mRNA sequence.  
ACCESSION AV629488.1 GI:10792122  
VERSION    AV629488  
KEYWORDS   EST.  
SOURCE     Chlamydomonas reinhardtii.  
ORGANISM   Chlamydomonas reinhardtii.  
REFERENCE  Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
AUTHORS    1 (bases 1 to 461)  
            Chlamydomonas reinhardtii.  
            Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,  
            Nakamura, Y. and Tabata, S.  
TITLE       Generation of expressed sequence tags from low-CO2 and high-CO2  
            adapted cells of Chlamydomonas reinhardtii  
JOURNAL     DNA Res. 7 (5), 305-307 (2000)  
MEDLINE     20539644  
COMMENT     Contact: Erika Asamizu  
            The First Laboratory for Plant Gene Research  
            Kazusa DNA Research Institute  
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
            Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.  
            1. .461  
            /organism="Chlamydomonas reinhardtii"
```

```
/strain="c9"  
/db_xref="taxon:3055"  
/clone="LCL059f09_r"  
/clone_lib="Chlamydomonas reinhardtii 5% to 0.04% CO2"  
/note="Vector: pBluescriptII SK-; Site.1: EcoRI; Site.2:  
XhoI; The cDNA library was constructed from cells cultured  
in a carbon stress acclimatized condition in which carbon  
dioxide concentration in the bubbling gas was changed from  
5% to 0.04%"  
  
BASE COUNT      73 a   156 c   143 g   89 t  
ORIGIN  
  
Query Match      34.8%; Score 457.8; DB 10; Length 461;  
Best Local Similarity 99.6%; Pred. No. 7.8e-79;  
Matches 459; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 8 ACATCGTATGCTGCTGAGTGGCCCTTGCTCAAGACGGCGGCTGGGCGGATG 67  
Db 1 ACATCGTATGCTGCTGAGTGGCCCTTGCTCAAGACGGCGGCTGGGCGGATG 60  
  
QY 68 TGACTGTTGGCTGCTTATTGAGCTGCTCAAGCGGGCCACCGGCTCATGACCATTTGCC 127  
Db 61 TGACTGTTGGCTGCTTATTGAGCTGCTCAAGCGGGCCACCGGCTCATGACCATTTGCC 120  
  
QY 128 CTCCTAGGACCATGACGCTGACGCTGGGACACCTCGGTGGTCTGTCGACATCATGGCG 187  
Db 121 CTCCTAGGACCATGACGCTGACGCTGGGACACCTCGGTGGTCTGTCGACATCATGGCG 180  
  
QY 188 AGAAGTCCGCTACTTCCACTCCATCAAGAAGGGCGTGCACCGGCTGTGGATTGACCAC 247  
Db 181 AGAAGTCCGCTACTTCCACTCCATCAAGAAGGGCGTGCACCGGCTGTGGATTGACCAC 240  
  
QY 248 CCTGTTCTCTGCGCAAGGCTCTGGGCAAGACCGGCTCCAAAGCTGTACGCGCCCGCTCCG 307  
Db 241 CCTGTTCTCTGCGCAAGGCTCTGGGCAAGACCGGCTCCAAAGCTGTACGCGCCCGCTCCG 300  
  
QY 308 GCGTGACTACTGACCAACCAACGCGCTTCGCCCTCTTCTGCAAGGCGGCTATTGAGG 367  
Db 301 GCGTGACTACTGACCAACCAACGCGCTTCGCCCTCTTCTGCAAGGCGGCTATTGAGG 360  
  
QY 368 CTGCGCGGCTGCTGCTTCGCCCGCCCGGAGGAGTGCCTTCTGTCGCCCAACGACTGGC 427  
Db 361 CTGCGCGGCTGCTGCTTCGCCCGCCCGGAGGAGTGCCTTCTGTCGCCCAACGACTGGC 420  
  
QY 428 ACTCGCGCTGCTGCTGCTTCCTGCTGAAGGAGGAGTACAG 468  
Db 421 ACTCGCGCTGCTGCTGCTTCCTGCTGAAGGAGGAGTACAG 461
```

```
RESULT 3  
AV622787  
LOCUS      AV622787 Chlamydomonas reinhardtii 5% to 0.04% CO2 Chlamydomonas  
DEFINITION reinhardtii cDNA clone LCL054e04_r 5', mRNA sequence.  
ACCESSION AV622787  
VERSION    AV622787.1 GI:10771964  
KEYWORDS   EST.  
SOURCE     Chlamydomonas reinhardtii.  
ORGANISM   Chlamydomonas reinhardtii.  
REFERENCE  Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
AUTHORS    1 (bases 1 to 458)  
            Chlamydomonas reinhardtii.  
            Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,  
            Nakamura, Y. and Tabata, S.  
TITLE       Generation of expressed sequence tags from low-CO2 and high-CO2  
            adapted cells of Chlamydomonas reinhardtii  
JOURNAL     DNA Res. 7 (5), 305-307 (2000)  
MEDLINE     20539644  
COMMENT     Contact: Erika Asamizu  
            The First Laboratory for Plant Gene Research  
            Kazusa DNA Research Institute  
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
            Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
```

[illegible]

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakamuekazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

BASE COUNT	80 a	170 c	159 g	101 t	1 others
ORIGIN					

Query Match	30.1%	Score	395.6;	DB	10;	Length	511;
Best local similarity	98.8%;	Pred.	No. 8.6e-67;				
Matches	398;	Conservative	0;	Mismatches	5;	Indels	0;
QY	1	CGCGTGGACATCGATGATGGTTCGTCTGAGTGCGCCCTTGGTCCAAAGCGGGGGCGCTG	60				
Db	109	CGCGTGGACATCGATGATGGTTCGTCTGAGTGCGCCCTTGGTCCAAAGCGGGGGTCTG	168				
QY	61	GGGATGTGACTGTGTGGCCCTGCCATTATGAGCTGTGAAGCGGGGCCACCGCGGTCAIGACC	120				
Db	169	GGCATGTGACTGTGTGGCCCTGCCATTATGAGCTGTGAAGCGGGGCCACCGCGGTCAIGACC	228				
QY	121	ATTGCCCCCTCGCTACGACACAGTACGCTGACGCCCTGGGACACCTCGGTGGTCTGTGGACATC	180				
Db	229	ATTGCCCCCTCGCTACGACACAGTACGCTGACGCCCTGGGACACCTCGGTGGTCTGTGGACATC	288				
QY	181	ATGGGGCAGAAAGGTCCCGCTACTTCCACTCCATCAAGAAAGGGCGTGCACCGCGTGTGGATT	240				
Db	289	ATGGGGCAGAAAGGTCCCGCTACTTCCACTCCATCAAGAAAGGGCGTGCACCGCGTGTGGATT	348				
QY	241	GACCAACCCCTGGTTCTCTGGGCCAAGGTCCTGGGGCAAGACCGCGTCCAAGCTGTACGGCCCC	300				
Db	349	GACCAACCCCTGGTTCTCTGGGCCAAGGTCCTGGGGCAAGACCGCGTCCAAGCTGTACGGCCCC	408				
QY	301	CGCTCCGGCGCTGACTACCTGGACAACCAACAGCGCTTCGCGCCCTTCTCTGAAAGGCCGT	360				
Db	409	CGCTCCGGCGCTGACTACCTGGACAACCAACAGCGCTTCGCGCCCTTCTCTGAAAGGCCGT	468				
QY	361	ATTGAGGCTGCCCGGTGCTGCCCTTCGGCCCCCGCGGAGGACT	403				
Db	469	ATAGAGGCTGCCCGGTGCTGCCCTTCGGCCCCCGCGGAGGACT	511				

RESULT	6
AY109531	
LOCUS	
DEFINITION	linear
ACCESSION	CL1198_1 mRNA sequence.
VERSION	AY109531.1 GI:21213285
KEYWORDS	HTC.
SOURCE	Zea mays.
ORGANISM	Zea mays. Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 2147) Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whittsett,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V. Maize Mapping Project/DuPont Consensus Sequences for Design of Oveργο Probes Unpublished (2002)
REFERENCE	2 (bases 1 to 2147)
AUTHORS	Coe,E.C.
TITLE	Direct Submission
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
FEATURES	Location/Qualifiers . .2147
source	

/organism="Zea mays"
 /db_xref="MaizeDB:630100"
 /db_xref="taxon:4577"
 /clone="Cl1198_1"
 /clone_lib="Maize Mapping Project/DuPont Cornsensus Library"
 /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

BASE COUNT	400 a	582 c	604 g	365 t	196 others
ORIGIN					
Query Match	29.3%; Score 384.6; DB 11; Length 2147;				
Best Local Similarity	58.0%; Pred. No. 1.5e-64;				
Matches 782; Conservative	0;	Mismatches 477;	Indels 90;	Gaps	5;
QY	5	TGGACATCGTATGGTGTCTGCTGAGTTCGGCCCTTGGTCCAAAGACGGGGCGCTGGGCG	64		
Db	230	TGAACGTGCTTCTGTCGGCGCCGAGATGGCGCGTGGAGCAAGACCGGGCGCTCGCG	289		
QY	65	ATGTGACTGGTGGCTTCGCTATTGAGCTGGTTCAGCGGGGGCCACCGGCTCATGACCAATG	124		
Db	290	ACGNNNNNNNNNNNNNNNNNNNNNATGGCGCGGAACGGGCAACGGTCTCATGTCTGCTCT	349		
QY	125	CCCTCTGCTAGCACGACGTACGCTGACGCTGGGACACCTCGGTGGTGACATCA---	181		
Db	350	CTCCCGCTACGACAGTACAAAGACGCTGGGACACCGAGTGTCTCCGAGATCAGA	409		
QY	182	-----TGGGGAGAAAGTTCGGCTACTTCCACTTCCATCCATCAAGAGGGCTGCACGCG	232		
Db	410	TGGAGACGGGTACGACGCGTCAAGCTTCTTCCACTGCTACAAAGCGGGAGTGCACGCG	469		
QY	233	TGTGATGTACACCCCTGGTTCTGTCGCCAAGGCTTGGGGCAAGCGGCTCCAAAGTGT	292		
Db	470	TGTTGTTTGAACCCCACTGTCTCTGGAGGGNNNNNNNNNNNNNNNNNNNNNNNNNNCT	529		
QY	293	ACGGCCCCGCTCGGGGCTGACTACCTTGGACACCAACCAAGCGCTGCGCCTGTTCTGCA	352		
Db	530	ACGGGCTGTCGTGGAACGGACTACAGGGACAACCAAGCTCGGCTTACGCTGCTATGCC	589		
QY	353	AGGGCGTATTGAGCTGCCGCGTGTGCGCCCTTCGCGC-----	390		
Db	590	AGGACGACATTGACGCTCCAAAGTACTTACGCTTCAACAAACCAACCACTTCTCCGGAC	649		
QY	391	--CCGGGGAGGACTCGCTCTGCTGGCCAAACGACTGGCACTCGCGCCTGCTGCGCGTCC	448		
Db	650	CATACGGGAGAGCTGCTGTGCTGCTGCAACGACTGCAACCGGCGCTCTCTCGTCT	709		
QY	449	TGCTGAAGGACGAGTACAGCCCAAGGCGCAGTTTCAACAAAGGCAAGTCGGTGTGGGTA	508		
Db	710	ACCTCAAGAGCAACTACCAGTCCCACGGCATCTACAGGGAGGCAAGACCGCTTCTGCA	769		
QY	509	TCCACATCGCTTCCAGGGCGCATGTGGAGGAGGCTTCAAGGACACGACGCTGC	568		
Db	770	TCCACATCTCTTACAGGGCGGTTGGCTTCCCGACTACCGGAGCTGAACCTCC	829		
QY	569	CCAGGCGCGCTTTGACAAGCTGGCCCTTCTCGGAGGCTATGCCAAGGTTTACACTGAGG	628		
Db	830	CGGAGAGATTCAAGTCGTCCTTCGATTTCATCGACGGGTACGAGAAG-----	876		
QY	629	CCACCCCATGGAGGAGACGAGAGCCCCGCTCAGGGAAAGACCTACAAAGATCA	688		
Db	877	-----CCGCTGGAAG-----	898		
QY	689	ACTGCTGAAGGGTGGCATATTACGCGCGCCACAAAGCTGGTGACTGTGTGCGCCCACTACG	748		
Db	899	ACTGGATGAAGCCGGGATCTCGAGGCGGACAGGCTCTCACCGTCAGCCCTACTACG	958		
QY	749	CGACGAGATCGCTGCCGATCGCGCGCGGTGTGAGCTGGACACCGTCTATCCGCGCA	808		

Db 959 CCGAGGAGCTCATCTCCGGCATCCAGGGGCTCGAGCTCGAACAATCATGCGCCCTCA 1018

QY 809 AGGCATTTAGGCGATTGTGAACGGCATGACATTTAGGATGGAAACCCCAAGACCGACA 868

Db 1019 CCGGCATCACCGGCATCTCAACGGCATGGAGCTCAGCGAGTGGACCCCAAGCGGACA 1078

QY 869 AGTTCTCTGTGGCCCTAGCAGCAGACAGAGCTCTACGCGGCAAGCCGCCCAAGG 928

Db 1079 AGTACATCGCGGTGAAGTACACGTGTCGAGCGCCGTGGAGGCGCAAGCGGTGAACAAG 1138

QY 929 AGGCCCTGACAGCGGAGCTGGGCTGCTGTGGACCCCAACCGCCCTCTGTTCGCCCTTCA 988

Db 1139 AGCGCTGACAGCGGAGCTCGGCTCCCGGTGGACCGGAACATCCCGTGGTGGCGTTCA 1198

QY 989 TCGGCCCTTGGAGGACAGAGGCTGGACATCATCTCTGGCGCGCCCTGCCCCAAGATCC 1048

Db 1199 TCGGCAGGCTGGAGAGACAGAGGCGCCCGACGCTCATGGCGCGCCGACATCCCGCAGCTCA 1258

QY 1049 TGGCCACCCCAAG- --GTGCAGATGCCATCTCTGGGTACCGCAAGCGCGCCTACGAGA 1105

Db 1259 TGGAGATGGTGGAGGAGTGCATATCTGCTGGGACCGGCAAGAGATTCGAGC 1318

QY 1106 AGCTGGTGAACGCATCGGCACCAAGTAAAGGCGCGCAAGGCGGTGGTCAAGTTCT 1165

Db 1319 GCATGCTCATGAGCGCGAGGAGAAGTTCGCCAGGCAAGTGGCGCGTGGTCAAGTTCA 1378

QY 1166 CGCGGCCCTTGGCGGCACATCTCACCGCGCGCGGCTGCTGCTGGTGGCGCTCGCGCT 1225

Db 1379 ACGGCGCTTGGCGCACCATATGCGCGCGCGCGAGCTGCTGCGCGTCAACGACCGCT 1438

QY 1226 TCGAGCCCTCGCGCTGATCAGCTGCACGCCCATCTACGCTACCGTACCGTGGCGTGTAG 1285

Db 1439 TCGAGCCCTCGCGCTCATCTACGCTCAGGGGATGCGATACGGAACCCCTGCGCCTGGC 1498

QY 1286 CTTCCACCGCGCGCTGGTGCAGACCGCTC 1314

Db 1499 CGTCCACCGGTGACTGCTGACACCATC 1527

RESULT 7
AW757933
LOCUS 874004E12.y1 C. reinhardtii CC-1690, Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence. EST 03-MAY-2000

DEFINITION
AW757933
VERSION
AW757933.1 GI:7687285
KEYWORDS
EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE 1 (bases 1 to 475)
AUTHORS Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P., McDermott,J.P., Silflow,C., Stern,D. and Surzycki,R.
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants; project phase 2
JOURNAL Unpublished (2000)
COMMENT Contact: Elizabeth H. Harris
DCMB Box 91000
Duke University
Durham, NC 27708-1000, USA
Tel: 919 613 8164
Fax: 919 613 8177
Email: chlamy@duke.edu.
Location/Qualifiers
1. .475
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Lambda Zap II"
/note="Vector: pBluescript II SK-; Site_1: ECORI; Site_2: XhoI; This library, constructed by John Davies and Jeffrey

McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (acetate-containing) medium in the light, TAP medium in the dark, HS (minimal) medium in ambient levels of CO₂ and HS medium bubbled with 5% CO₂. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with ExAssist (Stratagene) phage."

BASE COUNT 74 a 160 c 143 g 98 t
ORIGIN

Query Match 27.8%; Score 365.2; DB 10; Length 475;
Best Local Similarity 97.9%; Pred. No. 6.5e-61;
Matches 370; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 CGCTGGACATCGTATGTTGCTGCTGAGGTCGCCCTTGTGTCACAGACGGCGGCTGG 61

Db 98 CGCTGGACATCGTATGTTGCTGCTGAGGTCGCCCTTGTGTCACAGACGGCGGCTGG 157

QY 62 GCGATGTGACTGTGGCTGCTGCTATTGAGCTGGTCAAGCGCGCCACCGCGTCAATGACCA 121

Db 158 GCGATGTGACTGTGGCTGCTGCTATTGAGCTGGTCAAGCGCGCCACCGCGTCAATGACCA 217

QY 122 TTGCCCCCTGCTACGACAGTACGCTGACGCTGGGACACCTCGGTGGTGGACATCA 181

Db 218 TTGCCCCCTGCTACGACAGTACGCTGACGCTGGGACACCTCGGTGGGACATCA 277

QY 182 TGGCGGAGAGTCCGCTACTTCCACTCCATCAAGAGGGCGTGCACCGCTGTGGATTG 241

Db 278 TGGCGGAGAGTCCGCTACTTCCACTCCATCAAGAGGGCGTGCACCGCTGTGGATTG 337

QY 242 ACCACCCCTGTTCTTGGCCAAAGTCTGGGGCAAGACCGGCTCCAAAGCTGTACGCGCCCC 301

Db 338 ACCACCCCTGTTCTTGGCCAAAGTCTGGGGCAAGACCGGCTCCAAAGCTGTACGCGCCCC 397

QY 302 GTCGCGGCTGACTACTGTCGACCAACCAAGCGCTTGCCTCTGTCGAAGCGCGCTA 361

Db 398 GTCGCGGCTGACTACTGTCGACCAACCAAGCGCTTGCCTCTGTCGAAGCGCGCTA 457

QY 362 TTGAGGCTGCCCGGTGC 379

Db 458 TTGAGGCTGCCCGGTGC 475

RESULT 8
AW644765
LOCUS AV644765
DEFINITION AV644765 Chlamydomonas reinhardtii 5% CO₂ linear EST 15-DEC-2000
cDNA clone HCL093d12_r 5', mRNA sequence.
AV644765
VERSION AV644765.1 GI:10788093
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE 1 (bases 1 to 418)
AUTHORS Chlamydomonas reinhardtii.
TITLE Chlamydomonas reinhardtii.
JOURNAL Chlamydomonas reinhardtii.
MEDLINE Chlamydomonadaceae; Chlamydomonas.
COMMENT Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K., Nakamura,Y. and Tabata,S.
Generation of expressed sequence tags from low-CO₂ and high-CO₂ adapted cells of Chlamydomonas reinhardtii
DNA Res. 7 (5), 305-307 (2000)
20539644
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yanaka 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp; URL:http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1. 418
/organism="Chlamydomonas reinhardtii"


```
Query Match 21.2%; Score 278.4; DB 14; Length 702;
Best Local Similarity 66.3%; Pred. No. 4.3e-44;
Matches 399; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

QY 682 AAGATCAACTGGCTGAAGGCTGGCATATCGCCGCGACAGCTGGTGAAGTGTGCGCC 741
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 103 AAGATCAACTGGATGAAGCGCGGATCTCGAGGCCGCGACAGGTGCTGAGTGAAGCC 162
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 742 AACTACGCGACGAGATCGCTGCGGATGCGCGCGCGGCTGGAGCTGGACACCGTCATC 801
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 163 TACTACGCGGAGGAGCTCATCTCTGGGGAAGCCAGGGGCTGCGAGCTCGACACATCATG 222
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 802 CGGCGCAAGGCGATGAGGCGCATGTGAACGCGATGACATTTAGGAGTGAACCCCAAG 861
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 223 CGCTCACTGGGATCACCGCATCGTCAACGGCATGTAGCGAGTGGGACCCACC 282
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 862 ACGGACAAGTTCCTGTCTGCGCCCTACGACGACAAGCGCTACGCGCGCAAGCGCCGC 921
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 283 AAGGACAAGTTCCTGCGCGTCACTAGACATCACACCGCTTGGAGGGAAGCGCGTG 342
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 922 GCGAAGAGCGCTGAGCGCGAGCTGGGCTGCTGTGTGAACCCCGCCCGCTGTTTC 981
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 343 AACAAAGAGCGCTGAGCGCGAGTGGGCTGCGGCTGGACCGGAAGGTGCCCTGGTG 402
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 982 GCCTTCATCGCGCGCTGAGGAGGAGCAAGGGTGTGGACATCATCTTGGCGCGCCCTGCC 1041
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 403 GCGTTCATCGCGAGCTGGAGGAGCAAGGGGCGCGGACGATGATGCGCGCGCATCCCG 462
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1042 AAGATCCTGGCCACCCCAAGTSCAGATCGCATCTCGGTACCGGCAAGCGCGCTAC 1101
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 463 GAGATCCTGAAGGAGGAGCGTCCAGATCGTCTCTCTGCGCACCGGGAAGAGAAGTTC 522
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1102 GAGAAGCTGTGAAGCGCATCGGCACCAAGTACAAGGCGCGCGCCAGGCGGTGTCAG 1161
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 523 GAGCGGCTACTAAGAGCAATTTAGAGGAAATCCCGAGCAAGGTGAGGGCGGTGTCAGG 582
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1162 TTCTCGCGCGCCCTGGCGACATGCTCACCGCGCGCGCGCATCTGCTGCTGCTGCTG 1221
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 583 TTCAAGCGCGCTGCTCACCAGATGATGCGCGCGCGCGCGCATGCTGCTGCTGCTGCTG 642
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1222 CGCTTCAGCGCTGCGCGCTGATCCAGCTGACGCGCATGACGCTACGCTGCGCGCTG 1281
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 643 CGCTTCAGCGCTGCGCGCTCATCCAGCTCCAGGGATGCGCTACGGAAGCGCGTGGCG 702
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
AV628313
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT
FEATURES
source

AV628313 Chlamydomonas reinhardtii 560 bp mRNA linear EST 15-DEC-2000
reinhardtii cDNA clone LCL039h11_r 5', mRNA sequence.

AV628313.1 GI:10790947
EST.
Chlamydomonas reinhardtii.
Chlamydomonas reinhardtii.
Eukaryota: Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.

1 (bases 1 to 560)
Asanizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
Nakamura,Y., and Tabata,S.
Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
DNA Res. 7 (5), 305-307 (2000)
20539644
Contact: Erika Asanizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asanizukazusa.or.jp. URL: http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1..560
/organism="Chlamydomonas reinhardtii"
```

```
/strain="C9"
/db_xref="taxon:3055"
/clone="LCL039h11_r"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from cells cultured
in a carbon stress acclimatized condition in which carbon
dioxide concentration in the bubbling gas was changed from
5% to 0.04%"
BASE COUNT 92 a 186 c 173 g 109 t
ORIGIN

Query Match 20.8%; Score 272.8; DB 10; Length 560;
Best Local Similarity 99.3%; Pred. No. 5e-43;
Matches 274; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCGCTGACATCGTGGTTCCTGCTGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 285 GCGCTGACATCGTGGATGGTTCCTGCTGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTG 344
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 GCGATGTGACTGTGGCTGCCCTATTGAGCTGTCAAGCGCGCCACCGCGTCAATGACC 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 345 GCGATGTGACTGTGGCTGCCCTATTGAGCTGTCAAGCGCGCCACCGCGTCAATGACC 404
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 ATTGCCCTCGCTACGACAGTACGCTGAGCGCTGGGACACCTCGTGGTGGTGGACATC 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 405 ATTGCCCTCGCTACGACAGTACGCTGAGCGCTGGGACACCTCGTGGTGGTGGACATC 464
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 181 ATGGCGGAGAGGTCCTGCTGCTGCGCAAGTCTGGGGCAAG 276
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 465 ATGGCGGAGAGGTCCTGCTGCTGCGCAAGTCTGGGGCAAG 560
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 241 GACCACCCCTGCTTCCTGCGCAAGTCTGGGGCAAG 276
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 525 GACCACCCCTGCTTCCTGCGCAAGTCTGGGGCAAG 560
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
BE024926
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

BE024926 552 bp mRNA linear EST 06-JUN-2000
Chlamydomonas reinhardtii CC-1690, normalized, Lambda Zap II
Chlamydomonas reinhardtii cDNA, mRNA sequence.

BE024926.1 GI:8287367
EST.
Chlamydomonas reinhardtii.
Chlamydomonas reinhardtii.
Eukaryota: Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.

1 (bases 1 to 552)
Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,
McDermott,J.P., Silflow,C., Stern,D. and Surzycki,R.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants; project phase 2
Unpublished (2000)
Contact: Elizabeth H. Harris
DCMB Box 91000
Duke University
Durham, NC 27708-1000, USA
Tel: 919 613 8164
Fax: 919 613 8177
Email: chlamy@duke.edu.
Location/Qualifiers
1..552
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 2lgr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
II"
/note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; This library, constructed by John Davies and Jeffrey
McDermott, combines cDNAs from CC-1690 cells grown to
```

mid-log phase in TAP (acetate-containing) medium in the light, TAP medium in the dark, HS (minimal) medium in ambient levels of CO₂ and HS medium bubbled with 5% CO₂. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."

```

BASE COUNT      91 a  185 c  167 g  109 t
ORIGIN

Query Match      19.2%; Score 251.8; DB 10; Length 552;
Best Local Similarity 99.2%; Pred. No. 5.8e-39;
Matches 253; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGCCTGGACATCGTGATGTTGCTGCTGAGGTGCGCCCTTGGTCCCAAGACGGCGGCTG 60
Db 298 GGCCTGGACATCGTGATGTTGCTGCTGAGGTGCGCCCTTGGTCCCAAGACGGCGGCTG 357
QY 61 GCGGATGTGACTGTGCTGCTGCTATTGAGCTGGTCAAGCGCGCCACCGCGTCATGACC 120
Db 358 GCGGATGTGACTGTGCTGCTGCTATTGAGCTGGTCAAGCGCGCCACCGCGTCATGACC 417
QY 121 ATTGCCCTCGCTACGACAGTACGCTGACGCTGGGACACCTCGGTGGTGGGACATC 180
Db 418 ATTGCCCTCGCTACGACAGTACGCTGACGCTGGGACACCTCGGTGGTGGGACATC 477
QY 181 ATGGGCGAGAAGTCCGCTACTTCCACTCCATCAAGAAGGGCGTGCACCGCGTGGATT 240
Db 478 ATGGGCGAGAAGTCCGCTACTTCCACTCCATCAAGAAGGGCGTGCACCGCGTGGATT 537
QY 241 GACCACCCCTGGTTC 255
Db 538 GACCACCCCTGGTTC 552

RESULT 13
AV641583
LOCUS      AV641583      517 bp      mRNA      linear      EST 15-DEC-2000
DEFINITION Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
cDNA clone HCL036908_r 5', mRNA sequence.
ACCESSION AV641583
VERSION   AV641583.1 GI:10784911
KEYWORDS  EST.
SOURCE    Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 517)
AUTHORS   Asamizu,E., Miura,K., Kuchro,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
Nakamura,Y. and Tabata,S.
TITLE     Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
JOURNAL   DNA Res. 7 (5), 305-307 (2000)
MEDLINE   20539644
COMMENT   Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1. .517
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone="HCL036908_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"

BASE COUNT      91 a  185 c  167 g  109 t
ORIGIN

Query Match      19.2%; Score 251.8; DB 10; Length 552;
Best Local Similarity 99.2%; Pred. No. 5.8e-39;
Matches 253; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGCCTGGACATCGTGATGTTGCTGCTGAGGTGCGCCCTTGGTCCCAAGACGGCGGCTG 60
Db 298 GGCCTGGACATCGTGATGTTGCTGCTGAGGTGCGCCCTTGGTCCCAAGACGGCGGCTG 357
QY 61 GCGGATGTGACTGTGCTGCTGCTATTGAGCTGGTCAAGCGCGCCACCGCGTCATGACC 120
Db 358 GCGGATGTGACTGTGCTGCTGCTATTGAGCTGGTCAAGCGCGCCACCGCGTCATGACC 417
QY 121 ATTGCCCTCGCTACGACAGTACGCTGACGCTGGGACACCTCGGTGGTGGGACATC 180
Db 418 ATTGCCCTCGCTACGACAGTACGCTGACGCTGGGACACCTCGGTGGTGGGACATC 477
QY 181 ATGGGCGAGAAGTCCGCTACTTCCACTCCATCAAGAAGGGCGTGCACCGCGTGGATT 240
Db 478 ATGGGCGAGAAGTCCGCTACTTCCACTCCATCAAGAAGGGCGTGCACCGCGTGGATT 537
QY 241 GACCACCCCTGGTTC 255
Db 538 GACCACCCCTGGTTC 552

RESULT 14
BQ804991
LOCUS      BQ804991      726 bp      mRNA      linear      EST 31-JUL-2002
DEFINITION whe3561_E02_J032s Wheat developing grains cDNA library Triticum
aestivum cDNA clone whe3561_E02_J03, mRNA sequence.
ACCESSION BQ804991
VERSION   BQ804991.1 GI:22029200
KEYWORDS  EST.
SOURCE    bread wheat.
ORGANISM  Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 726)
Altenbach,S., Anderson,O.D., Chao,S., Chln,A., Close,T.J., Cronin
,K., Crossman,C., Fenton,R.D., Lazo,G.R., Pham,J., Rausch,C.J.,
Wilson,C. and Woo,J.
The structure and function of the expressed portion of the wheat
genomes - Developing grains cDNA library
Unpublished (2002)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: SK primer.
Location/Qualifiers
1. .726
/organism="Triticum aestivum"
/cultivar="Butte 86"
/db_xref="taxon:4565"
/clone="WHE3561_E02_J03"
/clone_lib="Wheat developing grains cDNA library"
/tissue_type="whole grains"
/dev_stage="3-44 days post anthesis seed"
/lab_host="E. coli SOLR"
/note="Vector: Lambda ZAP II, excised phagemid; Site_1:
EcoRI; Plants were grown under six following different
environmental regimes in greenhouse, Environment 1)
24oc/17oc day/night, well-watered, with post-anthesis
fertilizer, Environment 2) 24oc/17oc day/night,

```


This Page Blank (uspto)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2003, 09:49:44 ; Search time 31.6082 Seconds
(without alignments)
12748.991 Million cell updates/sec

Title: US-09-980-771A-6

Perfect score: 1314

Sequence: 1 gcgcgtgacatcgtgatggt.....gcggcctggtgcacacgtc 1314

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.*

1: /cgn2.6/ptodata/1/ina/5A.COMB.seq.*

2: /cgn2.6/ptodata/1/ina/5B.COMB.seq.*

3: /cgn2.6/ptodata/1/ina/6A.COMB.seq.*

4: /cgn2.6/ptodata/1/ina/6B.COMB.seq.*

5: /cgn2.6/ptodata/1/ina/PCTUS.COMB.seq.*

6: /cgn2.6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	400.6	30.5	2267	4	US-08-679-645-25
2	388.4	29.6	2542	3	US-08-941-445A-6
3	237.2	18.1	2067	4	US-09-388-743-21
4	236.8	18.0	2274	4	US-09-388-743-17
5	218.8	16.7	2176	4	US-09-388-743-13
6	216.8	16.5	2202	4	US-09-388-743-1
7	202.6	15.4	2825	4	US-09-196-390-5
8	199.6	15.2	2380	1	US-08-572-951-3
9	193.2	14.7	2097	3	US-08-941-445A-10
10	191.4	14.6	2248	4	US-09-345-214-20
11	191	14.5	1798	4	US-09-345-214-16
12	191	14.5	2019	4	US-09-345-214-15
13	151.2	11.5	2007	3	US-08-941-445A-8
14	151.2	11.5	2085	1	US-08-572-951-2
15	116.8	8.9	2239	4	US-09-196-390-1
16	109	8.3	2348	4	US-09-388-743-5
17	106.4	8.1	2418	4	US-09-388-743-25
18	103.6	7.9	1758	3	US-08-836-567-3
19	102.2	7.8	2360	3	US-08-836-567-9
20	101.4	7.7	1528	4	US-09-345-214-6
21	101.4	7.7	1620	3	US-08-941-445A-20
22	101.4	7.7	1752	3	US-08-941-445A-12
23	101.4	7.7	2008	4	US-09-345-214-12
24	101.4	7.7	2491	4	US-09-345-214-5
25	99.8	7.6	2383	4	US-09-192-909-1
26	97.4	7.4	4800	3	US-08-941-445A-4
27	96.6	7.4	2990	1	US-08-572-951-1

28	88.4	6.7	2793	3	US-08-836-567-7	Sequence 7, Appli
29	86.4	6.6	6027	2	US-08-968-542C-1	Sequence 1, Appli
30	83.6	6.4	1926	3	US-08-836-567-5	Sequence 5, Appli
31	82.6	6.3	1464	1	US-07-735-065-1	Sequence 1, Appli
32	82.6	6.3	1464	1	US-08-469-202-11	Sequence 11, Appli
33	82.6	6.3	1464	2	US-08-484-434C-11	Sequence 5, Appli
34	80.6	6.1	2277	1	US-08-676-967-5	Sequence 5, Appli
35	80.6	6.1	2277	2	US-09-098-487-5	Sequence 5, Appli
36	80.6	6.1	2277	2	US-08-836-567-11	Sequence 11, Appli
37	78.4	6.0	2549	1	US-08-470-720-2	Sequence 2, Appli
38	78.4	6.0	4964	1	US-08-470-720-5	Sequence 5, Appli
39	68.2	5.2	1601	3	US-08-735-491-1	Sequence 1, Appli
40	66.2	5.0	2303	3	US-08-836-567-1	Sequence 1, Appli
41	66.2	5.0	4168	3	US-08-836-567-11	Sequence 11, Appli
42	65.2	5.0	1722	4	US-09-385-028-15	Sequence 15, Appli
43	65.2	5.0	11604	4	US-09-385-028-13	Sequence 13, Appli
44	65.2	5.0	15079	4	US-09-385-028-1	Sequence 1, Appli
45	64.4	4.9	44377	2	US-08-804-227C-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-08-679-645-25
; Sequence 25, Application US/08679645
; Patent No. 6350934
; GENERAL INFORMATION:
; APPLICANT: Zwick, Michael G.
; APPLICANT: Edington, Brent E.
; APPLICANT: McSwiggen, James A.
; APPLICANT: Merlo, Patricia Ann Owens
; APPLICANT: Guo, Lining
; APPLICANT: Skout, Thomas A.
; APPLICANT: Young, Scott A.
; APPLICANT: Folkerts, Otto
; APPLICANT: Merlo, Donald J.
; TITLE OF INVENTION: COMPOSITION AND METHODS FOR
; TITLE OF INVENTION: MODULATION OF GENE EXPRESSION
; TITLE OF INVENTION: IN PLANTS
; NUMBER OF SEQUENCES: 1263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/679,645
; FILING DATE: July 12, 1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,135
; FILING DATE: July 13, 1995
; APPLICATION NUMBER: 08/300,726
; FILING DATE: September 2, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 219/247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 25:

```
;
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2267 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-679-645-25

Query Match          30.5%; Score 400.6; DB 4; Length 2267;
Best Local Similarity 60.08; Pred. No. 8.3e-70;
Matches 810; Conservative 0; Mismatches 449; Indels 90; Gaps 5;

QY 5 TGGACATCGTATGATGTTCTCTCTAGAGTCGCCCCCTTGTCCAAAGACGGCGGCCCTGGGCG 64
DB 418 TGAACGTCTCTCTCTCGCGCGCGAGATGCGCGCGTGGAGCAAGACCGCGGCCCTCGGCG 477

QY 65 ATGTGACTGTGGTGGCTGCTATTGAGCTGTGCTCAAGCGGGCCACCGGCTCATGACCAITG 124
DB 478 ACGTCTCTCGGGGGCGCTGCGCGCGCCATGCGCGCGCAATGAGGACCGGTGATGTCGTCT 537

QY 125 CCCCTCGCTAGACACAGTACGCTGACGCGCTGGGACACCTCGGTGCTGCTGGACAT----- 179
DB 538 CTCGCCGCTACGACAGTACAAGACAGCGCTGGGACACCGGTGCTGTCGGAGATCAAGA 597

QY 180 -----CATGGCGAGAGGTCCTCTACTTCCACTCCATCAAGAGGCGGTGACCGCG 232
DB 598 TGGGAGACAGGTACGAGACGGTCAAGTCTTCCACTGTCTCAAGCGCGGAGTGGACCGCG 657

QY 233 TGTGATTTGACCAACCCCTGTTCTGCGCAAGTCTGGGCAAGACCGGCTCCAAGCTGT 292
DB 658 TGTTCGTGTGACCAACCCACTGTTCTTCGTGAGAGGGTTTGGGAAAGACCGGAGGAGAATCT 717

QY 293 ACGGCGCCCGCTCGCGCGCTGACTACCTGGCAACCAACCAAGGCGTTCGCCCTGTTCTGCA 352
DB 718 ACGGCGCTGACCTGGAACGGACTACAGGACCAACAGCTGGGGTTCAGCCTGCTATGCC 777

QY 353 AGGCGCTATTGAGCTGCCCCGCTGCTGCCCTTCGGC----- 390
DB 778 AGGCGACATTGAAGCTCCAAGGATCCTGAGCTCAACAAACCAACCTATCTTCCTCGGAC 837

QY 391 ---CCGCGCGGAGCTGCTCTGCGGCAAGCTGCGACCTGCGCCCTGCTGCGCGTCC 448
DB 838 CATAGCGGGAGAGCTGCTGCTGCTGCTGCAAGCTGCGACACCGGCCCTCTCTGCTGT 897

QY 449 TGCTAAGAGCAGTACCAAGCCCAAGGCGCTTCCAAAGGCCAAGTGGTGGTGGCTA 508
DB 898 ACCTCAAGAGCAATACCAAGTCCCGGCTACAGGAGCAAGACCGGCTTTCTGCA 957

QY 509 TCCAAACATCCCTTCCAGGCGCGCTATGCGGAGGCGCTTCAAGGACAGAGCTGC 568
DB 958 TCCAAACATCTCTTCCAGGCGCGGCTTCCGCTTCTCCGACTACCGGAGCTGAAGCTCC 1017

QY 569 CCGAGCGCGCTTTGACAAGCTGGGCTTCTCGGAGCGGCTATGCCAAGGTTTACACTGAGG 628
DB 1018 CGGAGAGATTCAAGTCTGCTTCGATTTTCATCGACGCTACGAGAG----- 1064

QY 629 CCACCCCATGAGGAGGACGAGAGAGCCCGCTGACGGGAAAGACCTACAGAGATCA 688
DB 1065 -----CCCGTGAAG-----GCGGGAAGATCA 1086

QY 689 ACTGCTGAAGGTTGCAATTATCGCGCGGACAAAGCTGTGACTGTGTCGCCCAACTAG 748
DB 1087 ACTGGAAGGCGGGATCTCTCGAGGCGGACAGGGTCTCTACCGTCAAGCCCTACTAG 1146

QY 749 CGACCGATCGTGCCTGATGCGCGCGGCTGTGAGAGCTGACACCGCTATCCGCGCA 808
DB 1147 CCGAGAGCTCATCTCCGCGATCGCGAGGGCTGCGAGCTGCAACATCATGCGCCTCA 1206

QY 809 AGGGATTGAGGCGATTGTGACGGCATGACATTTGAGGAGTGGAAACCCACAGACCGCA 868
DB 1207 CCGGATCACCGGCTATCGTCAAGCGGATGAGCTCAGCGAGTGGGACCCAGCAGGAGCA 1266

QY 869 AGTTCGTCTGCGCCCTACGACCAAGAGCTCTACCGCGCAAGGCGCGCCCAAG 928
DB 928
```

RESULT 2

```
US-08-941-445A-6
; Sequence 6, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanning
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2542 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA to mRNA
```

```
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 453..2282
US-08-941-445A-6

Query Match      29.6%; Score 388.4; DB 3; Length 2542;
Best Local Similarity 59.1%; Pred. No. 2e-67;
Matches 795; Conservative 0; Mismatches 461; Indels 90; Gaps 4;

QY 5 TGGACATCGTATGTTGCTGCTAGCTGCGCCCTTGGTCAAGACGGGGGCGCTGGCG 64
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 700 TGAACGTCGTGTTGCTGGCGCCGAGATGGCCCTGGAGCAAGACCGCGCGCTCG 759
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 65 ATGTGACTGTGTGCTGCTATGAGTGTGTCAGAGCGGGCCGCGCTATGACCAT 124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 760 ACGTCCTCGTGGCTCCCTCCCTGCTGCTGCGATGGCCAGAGGTGATGTGATCT 819
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 125 CCCCTCGCTAGGACAGTACGCTGACGCTGGGACACCTGCTGGTGTGCTGACAT 184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 820 CTCCTCGTACGACAGCTACAGACGCTGGGATACAGGCTGTGGCTGAGATCAAG 879
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 185 -----GGAGAAGTCCGCTACTTCCACTCCATCAAGAAAGGGGTGACCGCG 232
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 880 TTGCAGACAGTACGAGAGGTGAGGTTTTCATGCTCAACGCTGGAGTGCACCGTG 939
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 233 TGTGGATTGACCCCTCGTTCCTGGCCAGAGTCTGGGGCAAGACCGGCTTCCAGG 292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 940 TGTTCATCGACCATCCGCTATCTCTGAGAGAGGTTTGGGNAACACCGTCAAGAT 999
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 293 ACGGCCCTCGCTCGGCGCTGACTACTGTGACACCAACAGCGCTTCGCCCTGTG 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1000 ACGGACCTGACATGGAGTTGATTAAGAAAGACACCAAGATGCTTTCAGCCT 1059
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 353 AGCGCGCTATTGAGCTGCGCGGTGCTGCCCTTCGCGCCGCTGCTGCGCCGCT 393
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1060 AGGACGACTCGAGGCTCTAGGATCTTAACCTCAACAAACACCCATACTTCAAG 1119
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 394 -----GGCGAGGACTGCTCTGTGGCCCAACGACTGGCACTCGCGCTGTGTC 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1120 CTTATGCTGAGGATGTTGTGCTGCTGCAACGACTGGCACTGGCCCTGCGGAGCT 1179
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 449 TGTGTAAGGACGAGTACACCCCAAGGGCCAGTTCCACAGGCGCAAGTGGTGTG 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1180 ACCTGAAGAACAACCTACAGCCCAATGGCATCTACAGGAATGCAAGGTTGCTT 1239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 509 TCCACAACATCGCTTCCAGGGCGCATGTGGGAGGAGGCTTCAAGACACAGAGT 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1240 TCCACAACATCTCTACAGGGCGGTTTCGCTTTCGAGGATTAACCTGAGCTGA 1299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 569 CCGAGGCGGCTTTGACAAGCTGCGCTTCTCGGACGCGCTATGCCAAGGTTTAC 628
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1300 CCGAGAGGTTTCAAGTCTATCTGATTTTCATCGACGGGTATGAC----- 1343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 629 CCACCCCATGAGGAGGACAGAGAGCCCGCTGACGGGAAGACCTACAAGAA 688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1344 -----ACGCGGGTGGAGGGGAGGAAGATCA 1368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 689 ACTGGCTGAAGGTTGACATATGCGCGCGACAGCTGGTGAATGCTGCGCCCACT 748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1369 ACTGGATGAAGCGCGGAACTCTGGAACCGCACAGGAGTGTACCGTGAAGCC 1428
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 749 CGACCGAGATCGCTGCGGATCGCGCGCGCTGTGTGAGCTGGACACCGCTCAT 808
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1429 CCGAGGAGCTCATCTCGCGCATCGCCAGGAGTCCGAGCTCGCAACATCATG 1488
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 809 AGGCATGAGGAGTGTGTAAGCGCATGGACATTTAGAGTGAAGACCCCAAGAC 868
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1489 CCGGCATCACCGGCTGCTCAAGGGCATGGACGTCAGCGAGTGGATCTTAGCA 1548
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 869 AGTTCCTGTGCGCCCTACGACCAAGACAGCGTCTACGCGGCAAGCGCGCC 928
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 1549 AGTACATCACGCCCAAGTACGACGCAACAGGCAATCGAGCGAAGCGCTGAACA 1608
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 929 AGGCCCTGAGCGCGAGCTGGCGCTGCTGTGGACCCACCGCCCTGTTCGCTTCA 988
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1609 AGCGCTTGCAGCGGAGCGGGTCTTCGGTGCAGAGAAATCCCTGATCGCGTTC 1668
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 989 TCGGCGCTGAGGAGCAGAGGCTGTGGACATCATCTGCGCCCTGCCCCAAGATCC 1048
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1669 TCGCAGGCTGAGGAAACAGAGGCGCTGACGTCATGCGCGCCCATCCCGAGCTCA 1728
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1049 TGGCCACCCCAAGGTGAGATCGCCATCTCTGGGTACCGGCAAGCCCGCTACG 1108
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1729 TG---CAGGAGGAGCTCCAGATGCTTCTCTGGTACTGGAAGAAGAGTTCGA 1785
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1109 TGTGAACGCCATCGGACCAAGTACAAGCGCGCGCAAGGCGGTGCTCAAGTTC 1168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1786 TGCTCAAGAGCATGAGGAGAGATATCCGGCAAGGTGAGGCGGTGTGAGTTCA 1845
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1169 CGCCCTGCGCACATGCTACCGCGCGCGCGACTTTCATGCTGTGCGCTCGCTTC 1228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1846 CGCGCTTGCTCATCTCATCATGCGCGGAGCGAGTGTGCGCGTCCCGAGCGCTTC 1905
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1228 AGCCCTGCGGCTGATCCAGCTGACGCCATGCTACTACGCTACCGTGGCTGTAG 1288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1906 AGCCCTGTGSACTATCCAGCTGAGGGATGAGATACGGAACGCCCTGTGCTT 1965
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1289 CCACCGCGGCGCTGCTGCACACCGTGC 1314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1966 CCACCGGTGGCTGCTGCACACCGTGC 1991
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 3

```
US-09-388-743-21
; Sequence 21, Application US/09388743
; Patent No. 6423886
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Ian
; TITLE OF INVENTION: No. 6423886el Starch Synthase Polynucleotides and Their
; TITLE OF INVENTION: Use in the Production of New Starches
; FILE REFERENCE: 1144
; CURRENT APPLICATION NUMBER: US/09/388,743
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 2067
; TYPE: DNA
; ORGANISM: Tulipa fosteriana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (58)...(1857)
US-09-388-743-21
```

```
Query Match      18.1%; Score 237.2; DB 4; Length 2067;
Best Local Similarity 52.0%; Pred. No. 5.3e-36;
Matches 702; Conservative 0; Mismatches 558; Indels 90; Gaps 4;
```

```
QY 1 CGCGTGCATCGTATGTTGCTGCTGAGTGTGCGCCCTTGGTCCAAAGACGGCGGCTG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 271 GGGATGAACCTGTGTTTCGTGGGACGAGACGGGTCTGACAAAGACCGCGGGTCT 330
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 GCGCATGTGACTGTGGCTGCTGCTTATTGAGCTGCAAGCGCGCCACCGGCTCATGAC 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 331 GGGATGTGCTAGGAGGTTACCCCGCTGCGCGGAGAGGAGCATCGGTGCTGTTG 390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 ATTGCCCTCGCTACGACCACTGCTGAGCGCTGGGACCTCGTGGTGTGAGACAT 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 391 GTCACCTCCGGGTACGATCAAGAGTATGATGGGACACAAACTGCTGTTGATGTC 450
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 ATGGGC-----GAGAAGTCCGTACTTCCACTCCATCAAGAGGGCGTGCAC 228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 451 AAAGTCGGGGATAAAATGAGAGCGGTTCGTTCTTCCACCTCCATTAAGAGGGGGTGGAT 510
QY 229 CCGGTGTGATGACACACCCCTGGTTCCTGGCCAAAGTCTGGGCAAGACCGGCTCCAAG 288
Db 511 AGGTGTTTATGATCACCCCTTGGTTCTTGAGAAGGTTTGGGGAAAACCGGTGGGAA 570
QY 289 CTGTACGGCCCGCTCCGGCGCTGACTACCTTGGCAACACCAAGGCTTCGCCCTGTC 348
Db 571 TTGTATGTCTCTTACTGGAAGTATGATGATAAAGAGTCAAGGTTTCACTCTCTG 630
QY 349 TGAAGGCGCTATTTGAGGCTGCCCGCTGCTGCTCCCTTCGGCCCT----- 393
Db 631 TGTACGCTGCTTTGGAGCTCAAGAGTCTAAAGTCTCAACACAGTAAATATTTTCT 690
QY 394 -----GGCGAGGACTGGCTTCCTGGCCAAAGTGGGCAAGTGGGCACTCCGCTGGTGC 444
Db 691 GGACCATATGTTGAAGATGCTGTTTTTATGGCAAGATGGCACACTGGACCTCTCCA 750
QY 445 GTCTGCTGAAGACGAGTACACGCCCAAGGGCCAGTTCCAAAGGCCAAGTGGTGGT 504
Db 751 TGTACTTGAAGAGTGTATATAATCAGAAGGATTTATGAGAGTCCCAAGGTTGCTTT 810
QY 505 GCTATCCACACATCGCTTCCAGGCGCCATGCTGGGAGGAGCTTTCAGGACACGAAG 564
Db 811 TGAATTCATATATGCAATACCAAGGAGATTCGCCCTTTCCTGATTTCTGCTTCTCAAC 870
QY 565 CTGCCCGAGCGCTTTTACAGAGCTGGCTTCTCGAGCGGTATGCCAAGGTTTACACT 624
Db 871 CTCCAGACACATTTAAATCTCGTTTATTTCTTCGATGGATATACAAACCT----- 924
QY 625 GAGGCCACCCCATGAGGAGGACGAGAGAGCCCGCTGACGGGAAGAGCTACAGAAG 684
Db 925 -----GTGAAGGTTAGAAAA 939
QY 685 ATCAACTGCTGAAGGTGSCATTATCGCGCGGACAAAGTGTGACTGTGCGCCCAAC 744
Db 940 ATAAATGATGAAGCTGGAATTTGGAAGCGGACACTGTTTAACTGTGAGCCCGTAT 999
QY 745 TAGCGACCGAGATCGCTCGGATCGCGCGGCTGTGAGCTGGACACCGTCATCCGC 804
Db 1000 TATGCTAAAGAGCTGCTCTGAGAGATAGAGGTGTGAGTGGACACGTTCTGCGC 1059
QY 805 GCAAGGGATGAGGCAATTTGACGCAATGAGGATGAGAGTGGAAACCCCAAGACC 864
Db 1060 TTGAGGGCGCTCAAGAGAAATTTGATGAGTGGATGATTAATGTGGAATCCATTGACA 1119
QY 865 GACAAGTTCCTGCTCGCCCTACGACCAAGACAGGCTCTACCGCGCAAGCGCGCGC 924
Db 1120 GACAAATTTACTGCAATTTACGATGCAACATGTTACAGAGCAAAACGTTTAAAT 1179
QY 925 AAGGAGCGCTGACGCGGAGCTGGGCTTGTGGACCCCAAGCGCCCGCTGTCGCC 984
Db 1180 AAGCAAGAATTACAAGCAGAAGTTGGCTTGCCTGTAGATCCAGACATTCCTGTTATAGTT 1239
QY 985 TTCAATCGCGCTGTGAGGACGAGAGGTTGGACATCATCTGCGCCCGCTGCCCAAG 1044
Db 1240 TTTGTTGGAAGCTGTGAGGACGAGAGGTTGAGATTTCTAGCTGCAAGAAATTCAGAA 1299
QY 1045 ATCTTGCCACCCCAAGTGTGACATCGCATCTCGGTACCGGCAAGCGCGCTACGAG 1104
Db 1300 TTAAGG---ATGAGAACGTTTCAATATTTCTTGAAGTGGCAAGAACACCTCGAA 1356
QY 1105 AAGCTGGTGAACCCATCGGCAACAGATACAGGGCGCGGCGGCTTCATGCTGGTCCCTCGC 1164
Db 1357 AAGGAGCTTGAAGAAATAGAAACAAATTTCCAGACAAGATGAGACTTGTGCGAAATTC 1416
QY 1165 TCGGCGCTGTGCGCATGCTCACGCGCGGCGGCTTCATGCTGGTGGCTCGCTCGC 1224
Db 1417 AATGTTCCGTTGGCTCATATGATGAGTGGCTGGAGGTTATTTTAAATATTTCTTAGTAGA 1476
QY 1225 TTGAGCCCTGCGGCTGATCCAGCTGCAGCGCATACGCTACCGTACCGTGGCGCGGTA 1284
Db 1477 TTTGAGCGGTGGCTTATTCAGCTTGAAGGCAAGAAATATGGGATGCCAGCCATGATG 1536

QY 1285 GCCTCCACCGCGCGCTGTGTCGACACGCTC 1314
Db 1537 TCACACCGGTGTCTTGTAGACAAATC 1566

RESULT 4

US-09-388-743-17
; Sequence 17, Application US/09388743
; Patent No. 6423886
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Lan
; TITLE OF INVENTION: No. 6423886el Starch Synthase Polynucleotides and Their
; FILE OF INVENTION: Use in the Production of New Starches
; FILE REFERENCE: 1144
; CURRENT APPLICATION NUMBER: US/09/388,743
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 2274
; TYPE: DNA
; ORGANISM: Typha latifolia
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (115)...(1956)
US-09-388-743-17

Query Match 18.0%; Score 236.8; DB 4; Length 2274;
Best Local Similarity 53.4%; Pred. No. 6.5e-38;
Matches 721; Conservative 0; Mismatches 537; Indels 92; Gaps 7;

QY 1 GCGTGGACATCGTGTGTTGCTGCTGAGTGCCTTGGTCCAAAGACGGCGGCTG 60
Db 373 GGATGACTATGCTTTTGGAGCTGAGATGCTCCATGGACAGACATGGAGGCTT 432
QY 61 GCGATGTGATGTTGGCTTGCCTTATGAGTGGTGAAGCGGCCACCGGCTCATGACC 120
Db 433 GGTGATCTTCTGGAGACTCCCGACCGCATTTGGCGCAATGACATCGAGTTATGTT 492
QY 121 ATTGCCCCCTGCTACGACAGTACGCTGACGCTGCGACACCTCGGTGGTGGACAT- 179
Db 493 ATAGCGCGCTATGATCAATACATGATGCTTGGATACAGATGCTTTGTTGAGTTG 552
QY 180 -----CATGGCGAGAAGTCCGCTACTTCCACTCCATCAAGAGGGCGTGCAC 228
Db 553 AAAGTTGGGATAGGTGTGAACCGTCCGCTCTTTCACCTGCTATAAAGAGGATGAT 612
QY 229 CGCGTGTGATTTGACACCCCTGGTCTTGGCCAAAGTCTGGGGCAAGACCGGCTCCAAG 288
Db 613 CGAGTTTTCGATACCCCTATGTTTCTTGGAAAGTCTGGGGGAAAACCTGGTGGGAA 672
QY 289 CTGTACGCGCCCGCTCCGGCGTGTGACTGCTGACAAACACACAGCGCTTCGCCCTGTC 348
Db 673 ATTATGGTCTTAACACTGGAACAGATATCAGGCAATCAGACTACGCTTACGTTCTTA 732
QY 349 TGAAGGCGCTATTTAGGCTGCGCGGTGCTGCTTCCGCGCC----- 393
Db 733 TGCAGGACGATTTGAAGCTCTTAGAATCTAAATCTCAACACAGTATCTTCTCT 792
QY 394 -----GGCGAGGACTGGGCTTCTGTCGCCAAAGCTGCGACTCGGCCCTGGTCC 444
Db 793 GGTCTTATGGGAAGATGTTATCTTCAATTTGCAATGATTGGCACACTTCTTCTGCCA 852
QY 445 GTCTGCTGAAGACGAGTACCGAGCCCAAGGCTTCCAGGAGGCAAGTTCGGTGTG 504
Db 853 TGCTACTTAAGAGCATGTACCATCCCGTGGCATTTACAGAACGCCAGGTTGCTTC 912
QY 505 GCTATCCACAAATCGCTTCCAGGGCGGATGAGGAGGCTTTCAAGAGCACAGGAG 564
Db 913 TGCATTCACAATATATCATACCAAGGTTCGATTTTCTCCCTCAGACTTCGAATTTCTCAAT 972

Db	1247	CATCGGGACCGCATGCCGTGGATCGGGGG---CAGGACGTGCAGCTGGTGGTATGCTGGG	1303
QY	1083	TACGGCAAGCGCGCCCTACGAGAAGCTGGTGAACGCATTCGCGACCAAGTACAAGGGCGG	1142
Db	1304	CACGGGCGCGACACTGGAACGAATGTCGAGCACTTGGAGCGGGAGCATCCCAACA	1363
QY	1143	CGCCAAAGGCGGTGCTCAAGTTCTCGGGGCGCCCTGGGGGCACATGCTACACGCGCGGGCGCGA	1202
Db	1364	GGTCGCGGGGTGGGTGGGKTTCTCGGTGCTATGGCGCATCGCATACGCGGGCGCGCA	1423
QY	1203	CTTCATGCTGGTGCCTTCGGCGCTTCAGAGCCCTCGCGGCGTATCCAGCTGCAGCGCATGCA	1264
Db	1424	CGTCTGGTATGCCCTCCCGCTTCAGAGCCCTCGCGGCTGAACCAAGCTTACCGCATGGC	1483
QY	1263	CTACGGTACCGTCGCGGTGGTAGCTTCCACGCGCGCGCTTGGTCGACACCGT	1313
Db	1484	ATACGGACACGCTCCCTCTGGTGGTCAGCCCGTGGCGGGCTTCAGGGACACGGT	1534

```

RESULT 13
US-08-941-445A-8
; Sequence 8, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 8:

```

Query Match 11.5%; Score 151.2; DB 3; Length 2007;
Best Local Similarity 57.3%; Pred. No. 2.9e-21;
Matches 361; Conservative 0; Mismatches 248; Indels 21;

RESULT 15
 US-09-196-390-1
 ; Sequence 1, Application US/09196390
 ; Patent No. 6307125
 ; GENERAL INFORMATION:
 ; APPLICANT: Block, Martina
 ; APPLICANT: Lorz, Horst
 ; APPLICANT: Lutticke, Stephanie
 ; APPLICANT: Walter, Lennart
 ; APPLICANT: Froberg, Claus
 ; APPLICANT: Kossmann, Jens
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
 ; TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH
 ; TITLE OF INVENTION: SYNTHESIS
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: James F. Haley, Jr., C/o Fish & Neave
 ; STREET: 1251 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States of America
 ; ZIP: 10020
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/196,390
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: DE 196 21 588.9
 ; FILING DATE: 29-MAY-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: DE 196 36 917.7
 ; FILING DATE: 11-SEP-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/EP97/02793
 ; FILING DATE: 28-MAY-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Haley, Jr., James F.
 ; REGISTRATION NUMBER: 27,794
 ; REFERENCE/DOCKET NUMBER: AGREVO-9
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 596-9000
 ; TELEFAX: (212) 596-9090
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2239 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Triticum aestivum L.
 ; STRAIN: cv. Florida
 ; HAPLOTYPE: ca. 21 d Caryopses
 ; IMMEDIATE SOURCE:
 ; LIBRARY: cDNA library in pBluescript sk (-)
 ; CLONE: TaSSS
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 3..2017
 ; US-09-196-390-1

Query Match 8.9%; Score 116.8; DB 4; Length 2239;
Best Local Similarity 52.4%; Pred. No. 1.5e-14;
Matches 331; Conservative 0; Mismatches 292; Indels 9;

Search completed: June 3, 2003, 15:31:39
Job time : 56.6082 secs

This Page Blank (uspto)

FILING DATE: September 2, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 219/247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 2267 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 25:

US-09-961-077-25
Query Match 30.5%; Score 400.6; DB 9; Length 2267;
Best Local Similarity 60.0%; Pred. No. 1.5e-91;
Matches 810; Conservative 0; Mismatches 449; Indels 90; Gaps 5;
QY 5 TGGACATCGTATGGTTCTGCTGAGGTGCGCCCTTGTGTCACAGACGGCGCCCTGGCGG 64
DB 418 TGAACGCTGCTTCTGTCGGCGCCGAGATGGCGCGTGGAGCAAGACCGCGGCTCGCGG 477
QY 65 AITGACGTGGTGGCTGCTATTGAGTGTGTCAGCGGGCCACCGCGTCATGACCAATTG 124
DB 478 AGTCTCTGGGGCTGCGCGGCCATGGCGGCAATGGGACCGTGTCAATGGTCTGCT 537
QY 125 CCGCTCGCTAGCAGCAGTACGCTGACGCTGGGACACCTCGGTGCTGCTGGACAT---- 179
DB 538 CTCCCGCTAGCAGCAGTACAGGACGCTGGACACCGCTGCTGCGAGATCAAGA 597
QY 180 -----CATGGCGGAGAGTCCGCTACTTCCATCCATCAGAGGCGCTGCACCGG 232
DB 598 TGGGAGACAGTACAGAGCGTACAGGTTCTTCCATGCTACAGCGCGAGTGGACCGG 657
QY 233 TGTGATTGACACCCCTGTTCTCTGGCCAGGTCTGGGGCAAGACCGGCTCCAAGCTGT 292
DB 658 TGTTCGTTGACCCACCTGTTCTTGGAGAGGTTTGGGAAAGACCGAGAGATCT 717
QY 293 ACGGCCCCGCTCGCGGCTGACTACCTGAGCAACCAAGGCTTGCCTGTTGTCGA 352
DB 718 ACGGCGCTGAGCGTGAACGAGTACAGGACACAGCTGCGGTGACCTGCTATGCC 777
QY 353 AGGCGCTATTGAGGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 390
DB 778 AGGCGAGCTTGAAGCTCAAGGATCCCTGAGGCTCAACAAACCCATCTCTCGGAC 837
QY 391 --CCGCGGAGAGTGCCTCTCTGTCGCAAGCTGAGCTGCGCTTCCGCTGTCGCTGCC 448
DB 838 CATACGAGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897
QY 449 TCTGAGGACGAGTACAGCCAGGCGGCTTCAAGGCGGCTTCAAGGCGGCTGCTGCTG 508
DB 898 ACCTCAAGAGCACTACAGTCCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
QY 509 TCCACATGCTGCTTCCAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 568
DB 958 TCCACATGCTTCCAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017
QY 569 CCGAGCGGCTTTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 628
DB 1018 CGGAGATTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1064
QY 629 CCACCCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 688
DB 1065 -----CCCGTGGAG-----GCCGAGATCA 1086
QY 689 ACTGGCTGAAGGTGCAATTATCGCGCGGACAAAGTGTGCTGCTGCTGCTGCTGCTGCTG 748
DB 1087 ACTGGATGAGCGCGGATCTCTGAGGCGGACAGGCTCTCACCCTGCTGCTGCTGCTGCTG 1146

QY 749 CGACGAGATCGTCCCGATGCCCGCGGCTGTGAGCTGAGACACCGTATCCCGGCA 808
DB 1147 CCGAGGAGCTATCTCCCGCATGCCCGGCTGCGAGCTGACAACTATCATCGCTCA 1206
QY 809 AGGCGATTGAGGCGATTGTGACGCGATGGACATTGAGAGTGGAAACCCCAAGACCGCA 868
DB 1207 CCGGATCACCGGCTATGTCACGCGATGGAGTCAAGGCTGAGGAGTGGACCCCAAGGAGCA 1266
QY 869 AGTCTCTGCTGCGCCCTAGCAGCAAGACAGCTGTACGCGCGGCAAGCGCGCGCAAGG 928
DB 1267 AGTACATCGCGTGAAGTACGAGCTGTGACCGCGCTGGAGCGCAAGGCGCTGAACAAGG 1326
QY 929 AGGCGCTGACGCGGAGTGGGCTGCTGTCGACCCACCGCCCGCTGCTGCTGCTGCTGCT 988
DB 1327 AGGCGCTGACGCGGAGTGGGCTGCTGTCGACCGGAGCAATCCCGTGGTGGGTTCA 1386
QY 989 TCGCGCGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1048
DB 1387 TCGGAGGCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1446
QY 1049 TGGCCACCCCAAG---GTGACATCGCCATCTGCTGACCGGCAAGCGCGCTACGAGA 1105
DB 1447 TGGAGATGGTGGAGACGTGACATGCTGCTGCGCACGGGCAAGAGAGTTCGAGC 1506
QY 1106 AGCTGCTGAACGCCATCGGCACCAAGTACAGGCGCGCGCCCAAGGCGTGGTCAAGTTCT 1165
DB 1507 GCATCTCATGACGCGCGAGGAGAGTTCACGAGCAAGTGGCGCGCTGCTCAAGTTCA 1566
QY 1166 CGCGCGCTGCGGACATGCTACCGCGCGCGGACATCATGCTGCTGCTGCTGCTGCTGCT 1225
DB 1567 ACGCGCGCTGCGGACCATCATGCGCGCGCGGAGAGTGTGCTGCTGCTGCTGCTGCTGCT 1626
QY 1226 TCGAGCGCTGCGGCTGATGACGCTGACCGCATGACGCTGACGCTGACGCTGACGCTGAC 1285
DB 1627 TCGAGCGCTGCGGCTGATGACGCTGACCGCATGACGCTGACGCTGACGCTGACGCTGAC 1686
QY 1286 COTCCACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1314
DB 1687 COTCCACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1715

RESULT 2

US-10-138-075-1
; Sequence 1, Application US/10138075
; Publication No. US20030087369A1
; GENERAL INFORMATION:
; APPLICANT: Broglie, Karen E.
; APPLICANT: Butler, Karlene H.
; APPLICANT: Harwell, Leslie T.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Orozco, Emil M.
; TITLE OF INVENTION: Granule-Bound Starch Synthesis
; FILE REFERENCE: BB1474 NA
; CURRENT APPLICATION NUMBER: US/10/138,075
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/288,315
; PRIOR FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 2399
; TYPE: DNA
; ORGANISM: Zea mays
US-10-138-075-1

Query Match 18.7%; Score 245.6; DB 9; Length 2399;
Best Local Similarity 52.9%; Pred. No. 1.3e-52;
Matches 715; Conservative 0; Mismatches 544; Indels 93; Gaps 5;
QY 1 GCGTGGACATCGTATGTTGCTGCTGAGTGGCGCCCTGTTGCTGCTGCTGCTGCTGCTGCTG 60
DB 503 GCGTGGACATCGTATGTTGCTGCTGAGTGGCGCCCTGTTGCTGCTGCTGCTGCTGCTGCTG 562

RESULT 9

US-09-894-633A-83
; Sequence 83, Application US/09894633A
; Patent No. US20020124285A1

; GENERAL INFORMATION:

; APPLICANT: Conner, Timothy
; APPLICANT: Dubois, Patrice
; APPLICANT: Malven, Marianne
; APPLICANT: Masucci, James
; TITLE OF INVENTION: PLANT REGULATORY SEQUENCES FOR SELECTIVE CONTROL OF GENE EXPRESSION
; FILE REFERENCE: 38-21(15856)B
; CURRENT APPLICATION NUMBER: US/09/894,633A
; CURRENT FILING DATE: 2001-06-28
; PRIORITY FILING DATE: 60/214,357
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 09/894,633
; PRIOR FILING DATE: 2000-06-28

; NUMBER OF SEQ ID NOS: 111

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 83

; LENGTH: 824

; TYPE: DNA

; ORGANISM: Zea mays

US-09-894-633A-83

Query Match 5.5%; Score 72.2; DB 10; Length 824;

Best Local Similarity 47.6%; Pred. No. 5e-09;

Matches 246; Conservative 0; Mismatches 268; Indels 3; Gaps 1;

```
QY 795 CGTCATCCGCGCCGAAGGCGATTTGAGGCGCATTTGTAACGGCATGGACATGAGGAGTGA 854
DB 132 CCGTCCGCGGCTTTCTTCCAGACTGCTTCGTACAGCGGGTTCGACGGTCTGGTCTGAT 191
QY 855 CCCCAGACGACAGATTCTCTGTCGCGCCCTACACAGACAGACGCTCTACGCGGCAA 914
DB 192 CGGTCCACCCAGTTCGAGAGTCGGAGCAGCAGCGGAGATCAACCACTCCCTCCCGG 251
QY 915 GSCCGCCGCCAAGGAGGCGCTGCAGGCGGAGTGGGCGCTTGCCTGTGGACCCGCGCC 974
DB 252 GGACGCGCTTCGACGCGGTTGGTGCAGCGCAAGTGGCGCTGGAGTGGCGCGGCT 311
QY 975 CCGTGTCCGCTTCATCGGCGCGCTGGAGAGCAGAGGTTGTGGACATCATCTGCGCGC 1034
DB 312 GGTGT---CCTGCGCGCATCTCTCGCATGCGCGTGGCGGCTGCTGATTAACCATACCG 368
QY 1035 CCGTCCCAAGATCCTTGGCCACCCCAAGGTGCAGATCGCATCTCTGGGTACCGCAAGC 1094
DB 369 CGGGCCCGGTTACCCGGTTCCGCTGGGGCGGAGGACTCGCTGCTGCTCGCCACGCG 428
QY 1095 CGCCTACGAGAGCTGGTGAACGGCATCGGCACCAAGTACAAAGGGCGCGCCGAAGGCGT 1154
DB 429 GCGCGAGCTGGAGCTGCGCGACGCGCAACTTCACCGTGGACCGCTCATCCAGATGTCG 488
QY 1155 GGTCAATTCCTGGCGCGCTGGCGACATGCTACCGCGCGCGGAGCTTATGCTGT 1214
DB 489 CGCCAAGGGGTTACGGGTGCAGAGAGTGGTGGCGCTGTCGCGCGCCACACGCTGGGTT 548
QY 1215 GCGCTCGCGCTTCGAGCGCTGCGGCTGATCCAGCTGCACGCGATGCACTAGCGTACC 1274
DB 549 CTCCCACTGCAAGGAGTTGCGCGACCGCGCTCTACACTTCCGAACAGGCGGGAGCC 508
QY 1275 GCGCGTGGTAGCTCCACCGCGGCGCTGGTTCGACACC 1311
DB 609 GGAGCAGTTCGACCCCGCAGCATGAACCGTCTTACGCC 645
```

RESULT 10

US-09-934-900-15

; Sequence 15, Application US/09934900

; Publication No. US20030054521A1

; GENERAL INFORMATION:

; APPLICANT: Booth, Russ

; APPLICANT: Cahoon, Rebecca E

; APPLICANT: Hitz, William D

; APPLICANT: Kinney, Anthony

; APPLICANT: Yadav, Naren

; TITLE OF INVENTION: Nucleotide Sequences of a New Class of Diverged Delta-9 Stearo

; TITLE OF INVENTION: ACP Desaturase

; FILE REFERENCE: BB1476 US NA

; CURRENT APPLICATION NUMBER: US/09/934,900

; CURRENT FILING DATE: 2001-08-22

; PRIOR APPLICATION NUMBER: 60/226996

; PRIOR FILING DATE: 2000-08-22

; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 15

; LENGTH: 1318

; TYPE: DNA

; ORGANISM: Oryza sativa

US-09-934-900-15

Query Match 5.3%; Score 70.2; DB 9; Length 1318;

Best Local Similarity 45.1%; Pred. No. 1.6e-08;

Matches 261; Conservative 0; Mismatches 318; Indels 0; Gaps 0;

```
QY 713 CGCGCGACAAGCTGGTGTGCTGCGCCCAACTACGCGACCGAGATCGCTGCGGATGCG 772
DB 286 CGGACTTCTCGCGGACTCGTCTCGGAGATGTTGAGCACCAGGTCACAGACTCCGCG 345
QY 773 CGCGGGGTGTGAGCTGAGACACCGTCATCCGGGCCAAGGCGATTGAGGCGCATTTGAACG 832
DB 345 CGCGCGCGCGGGCTCCCGCGAGAGTACTTCTGCTGCTGCTGGTGGGGACATGATTACCG 405
QY 833 GCATGACATTGAGGATGGAACCCCAAGACGACAAATTCCTGCTGCGCCCTAGCACC 892
DB 406 AGGAGCGCTGTCGACGATACAGACCATGATCAACACGCTGACGCGCTCCGCGAGAGA 465
QY 893 AGAACAGCGTCTACGCGCGCAAGGCGCGCCGAAGGAGCGCTGACAGCGCGAGCTGGCG 952
DB 466 CGCGGCCAGCGCTGCCCTCGGCGCTGTGACGCGGACCTGGAGCCCGGAGGAAACC 525
QY 953 TGCTGTGACCCCGCGCGCGCGCTGTGCGCTTCTATCGCGCGCTTGGAGGAGGAGG 1012
DB 526 GCCAGCGGACATCTCTCGGCAAGTACATGTACTCTCCGCGCGCTCGACATGCGCATGG 585
QY 1013 GTGTGACATATCTGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1072
DB 586 TCGAGAAGACCGCTCCAGTACTCATCGCGTCCGCGCATGATCGGGGAGCGGAGAACACC 645
QY 1073 CCATCTCTGTACCGCAAGGCGCTACGAGAGCTGTGTAACGCGCATCGGCAACCAAGT 1132
DB 646 CGTACCTGGGTTCTGTGTAACAGCTTCCAGAGCGCGGCGCGCGCGCTGTCGACGCGGA 705
QY 1133 ACAAGGCGCGCGCAAGGCGGTGTTCAAGTTCTCGCGCGCGCTTGGCGGACATGCTCACCG 1192
DB 706 ACAGCGCGCGCTCGCGAGGCGCGACGGGAGTACGCTCTCGCGCGCGCGCGCGCGCGCA 765
QY 1193 CGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1252
DB 766 TCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 825
QY 1253 AGCCCATGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1291
DB 826 GCGTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 864
```

RESULT 11

US-09-294-093B-3520

; Sequence 3520, Application US/09294093B

; Patent No. US20010051335A1

; GENERAL INFORMATION:

; APPLICANT: Lalgudi, Raghunath, V.

; APPLICANT: Ito, Laura, Y.

; APPLICANT: Sherman, Bradley, K.

; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL

; FILE REFERENCE: PL-0009 US

```
; CURRENT APPLICATION NUMBER: US/09/294, 093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 3520
; LENGTH: 299
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700380278H1
; LOCATION: 118
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-3520

Query Match          5.3%; Score 69.4; DB 10; Length 299;
Best Local Similarity 60.9%; Pred. No. 2.4e-08;
Matches 112; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1130 AGTAAAGGCGCGCAAGCGGTGCTCAAGTCTCGGCGCCCTGGCGCCACATGCTCA 1189
Db 75 AATACCATGCGCGAGTGAAGCTATGCTTAACCTATGAGGAGCCNCCTGTCACATTGATAT 134
QY 1190 CCSCGCGCGCGACTTCATGCTGGTGGCTCGGCTTCGAGCCCTGGGGCCTGATCCAGC 1249
Db 135 ATGCTGGCGGAGACTCACTTCCTGCTTCCTTCCTTCGAACTTGTGTCTAACACAGC 194
QY 1250 TGCACGCCATGACACTACGCTACCGTGGCTGGCTAGCTCCACCGCGGCGCTGTCGACA 1309
Db 195 TTATTCCTATGCGCTATGATCCATCCAGTTGTCGAAACTGGAGGCGCTGTACGACA 254
QY 1310 CCGT 1313
Db 255 CCGT 258

RESULT 12
US-09-966-881-40
; Sequence 40, Application US/09966881
; Patent No. US20020120960A1
; GENERAL INFORMATION:
; APPLICANT: Seymour, Graham
; Bird, Colin
; Medina-Suarez, Rosybel
; TITLE OF INVENTION: Genetic control Of Fruit Ripening
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Zeneca Ag Products Inc.
; STREET: 1800 Concord pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: USA
; ZIP: 19850
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/966, 881
; FILING DATE: 28-Sep-2001
; CLASSIFICATION: <Unknown>
; APPLICATION DATA:
; APPLICATION NUMBER: US/09/242, 860
; FILING DATE: 29-Mar-1999
; APPLICATION NUMBER: GB 9618862.8
; FILING DATE: 10-SEP-1996
; APPLICATION NUMBER: GB 9708366.1
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: PCT/GB97/02424
```

```
; FILING DATE: 08-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hohenschutz, Liza D.
; REGISTRATION NUMBER: 33,712
; REFERENCE/DOCKET NUMBER: SEE 50183/UST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 886-1699
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 804 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: U-D4
; SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-09-966-881-40

Query Match          5.2%; Score 68; DB 10; Length 804;
Best Local Similarity 51.6%; Pred. No. 5.7e-08;
Matches 195; Conservative 0; Mismatches 169; Indels 14; Gaps 2;

QY 1 GCGCTGCACATCGTGATGGTGTGCTGCTGAGTGCCTCCCTTGGTCCAAGA-CGGCGCGCCT 59
Db 349 GGGATGAACATAGTCTTTGTGCGCGCTGANAATGCTCCCTGGAGTAAACCGGANGACT 408
QY 60 GGGCGATGTGACTGGTGGCTGCTATTTAGCTGGTCAAGCGCGCGCCACCGCGTCATGAC 119
Db 409 CGGTGATGTTCTTGGAGGACTGCCACGCGCATGGCTGCGAAATGGACACAGANTGATGAC 468
QY 120 CATTCGCCCTCGCTACGACGAGTACGCTGACGCTGGGACACCTCGGTGGTTCG----- 172
Db 469 TATAGTCCACCCCTATCATCAGTACAAGGATGGGTGGGATACAAATGTCTCTGGCTGAAT 528
QY 173 -----TGGACATCATGGCGGAGAGGTTCGCTACTTCCACTCCATCAAGAGGCGCTGC 226
Db 529 AAAAGTTGGAATGAAATGAAANAATCCGCTTTCTTCCACTGCTATAAAAAAAGAAATG 588
QY 227 ACCGCGTGTGGATTGACACCCCTGTTCTCGCAAGGTCTGGGCAAGACCGCGCTCCA 286
Db 589 ACAGGTTTTCATGATCACCCTTTGTTCTTGAANAAGTGTGGGAAAAAACTGGAAGAAT 648
QY 287 AGCTGTACGGCCCCCGCTCGGCGCTGACTACCTTGGACACCAACCAACGCTTCGCCCTGT 346
Db 649 GAAATTTGTCCTNTCCCGGAACCTGATTTTNAAAAAACANCCCTTAAANTTTANCCCTT 708
QY 347 TCTCCAAGCGCGTATTG 364
Db 709 TTNCCGCGCCNCTTTTG 726

RESULT 13
US-09-934-900-11
; Sequence 11, Application US/09934900
; Publication No. US20030054521A1
; GENERAL INFORMATION:
; APPLICANT: Booth, Russ
; APPLICANT: Cahoon, Rebecca E
; APPLICANT: Hitz, William D
; APPLICANT: Kinney, Anthony
; APPLICANT: Yadav, Naren
; TITLE OF INVENTION: Nucleotide Sequences of a New Class of Diverged Delta-9 Stearic
; FILE OF INVENTION: ACP Desaturase
; FILE REFERENCE: BB1476 US NA
; CURRENT APPLICATION NUMBER: US/09/934, 900
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/226996
; PRIOR FILING DATE: 2000-08-22
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 1415
```

TYPE: DNA
ORGANISM: Zea mays
US-09-934-900-11

Query Match 5.2%; Score 68; DB 9; Length 1415;

Best Local Similarity 44.8%; Pred. No. 5.8e-08;

Matches 260; Conservative 0; Mismatches 320; Indels 0; Gaps 0;

```
QY 712 GCGCGGACAAAGCTGGTACTGTGCGCCCACTACGCGACGAGATCGCTCGGATGCC 771
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 346 GGGGACTTCTCCGGACTCTCTGTCGAGATGTCGCGACACAGTCCGCGAGCTGCCG 405

QY 772 GCGCGGCGTGTGAGTGACACCGCTCATCCGCGCCAAAGGCAITGAGGCGCATGTGTAAC 831
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 406 GCGCGCGCGCGGGCTCCCGACGAGTACTTCTGCTGCTGCTGCGGACATGCTACG 465

QY 832 GGCATGAGCATTTGAGGAGTGGAAACCCCAAGACCGCAAGTTCTGTCTCGCCCTACGAC 891
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 466 GAAGAGCGCTGCCACGCTACGAGACCATGATCAACACGCTGACGCGCTCCGCGACGAG 525

QY 892 CAGAACAGCGTCTACGCCGGAAGCGCGCCCAAGGAGGCGCTGCAGGCGCGAGCTGGGC 951
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 526 ACGGCGCCAGCAACTGCGCCCTGGCGGTCTGGACCGCGCTGGACATGCGCATG 585

QY 952 CTGCTGTGGACCCCAACCGCCCGCTTTGCGCTTCATCGCGCGCTGGAGGAGCAAG 1011
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 586 GCGCAGCGGACATCTCGCGCAAGTACATGTACTATCCGCGCGCTGACATGCGCATG 645

QY 1012 GGTGTGACATCATCTGCGCGCCGCTGCGCCAGATCTGCGCACCCCAAGTGCAGATC 1071
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 646 GTGAGAGACCGTCCAGTACCTCATCGGCTCGGATGATCCCGGAACGGAGAACAC 705

QY 1072 GCATCTGGGTACCGCAAGCGCGCTACGAGAGCTGGTGAACCCCATCGSCACCAAG 1131
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 706 CGGTACTGGCTTCTGTGTACAGAGCTTCCAGGACGCGCGCGCTCTCGCACGGC 765

QY 1132 TACAAGGCGCGCGCCAAAGCGGTGPTCAAGTTCTCGCGCGCCCTGCGCACATGCTACC 1191
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 766 AACACCGCGCGGTCCGCGAGCGCACGGGACGAGTCTCTGCGCGCGCTGCGGCACC 825

QY 1192 GCGCGCGGACTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1251
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 826 ATCGCGCGCGGAGAGCGGACGAGCGGCTGCGGCGCATCTGCGAGCGAGCTGCTG 885

QY 1252 CAGCGCATCACTACGCTACCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 886 CAGCTGGACCGGAGGCGCGGCTGCTGCGCGTCTGCGCGGACA 925
```

RESULT 14

US-10-138-221-6
Sequence 6, Application US/10138221
Publication No. US20020199216A1
GENERAL INFORMATION:
APPLICANT: Amy F. Macrae
TITLE OF INVENTION: USE OF TRANSPOSABLE ELEMENTS FOR ALTERING GENE EXPRESSION
FILE REFERENCE: 51178/7
CURRENT APPLICATION NUMBER: US/10/138,221
CURRENT FILING DATE: 2002-08-01
PRIOR APPLICATION NUMBER: US 60/287,882
PRIOR FILING DATE: 2001-05-01
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patent in version 3.1
SEQ ID NO 6
LENGTH: 4810
TYPE: DNA
ORGANISM: Zea mays
US-10-138-221-6

Query Match 5.1%; Score 66.8; DB 9; Length 4810;

Best Local Similarity 72.9%; Pred. No. 1.2e-07;

Matches 86; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```
QY 934 CTGACGGCGAGCTGGGCTGCTGTGGACCCACCGCCCTGTGCTTCCTTCATCGGC 993
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 CTGACGGCGAGCTGGGCTGCTGTGGACCCACCGCCCTGTGCTTCCTTCATCGGC 60

QY 994 CGCCTGGAGGACGAGGCTGTGGACATCATCTGCGCGCCCTGCCAAGATCTCG 1051
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 AGGCTGAAGAGCAGAGGCGCCCGACGTCATGCGCGCCCGCATCCCGACGTCATGG 118
```

RESULT 15

US-09-966-881-41

Sequence 41, Application US/09966881

Patent No. US20020120960A1

GENERAL INFORMATION:

APPLICANT: Seymour, Graham

Bird, Colin

Medina-Suarez, Rosybel

TITLE OF INVENTION: genetic control of Fruit Ripening

NUMBER OF SEQUENCES: 57

CORRESPONDENCE ADDRESS:

ADDRESSEE: Zeneca Ag Products Inc.

STREET: 1800 Concord Pike

CITY: Wilmington

STATE: DE

COUNTRY: USA

ZIP: 19850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/966,881

FILING DATE: 28-Sep-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/242,860

FILING DATE: 29-Mar-1999

APPLICATION NUMBER: GB 9618862.8

FILING DATE: 10-SEP-1996

APPLICATION NUMBER: GB 9708366.1

FILING DATE: 25-APR-1997

APPLICATION NUMBER: PCT/GB97/02424

FILING DATE: 08-SEP-1997

ATTORNEY/AGENT INFORMATION:

NAME: Hohenschutz, Liza D.

REGISTRATION NUMBER: 33,712

REFERENCE/DOCKET NUMBER: SEE 50183/UST

TELECOMMUNICATION INFORMATION:

TELEPHONE: (302) 886-1699

INFORMATION FOR SEQ ID NO: 41:

SEQUENCE CHARACTERISTICS:

LENGTH: 826 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

IMMEDIATE SOURCE:

CLONE: U-D10

SEQUENCE DESCRIPTION: SEQ ID NO: 41:

US-09-966-881-41

Query Match 5.1%; Score 66.6; DB 10; Length 826;

Best Local Similarity 55.9%; Pred. No. 1.3e-07;

Matches 185; Conservative 0; Mismatches 132; Indels 14; Gaps 3;

```
QY 1 GCGCTGACATCGTGTGCTGCTGAGTGGCTGCGCCCTGCTCCAGACGGCGGCTG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 307 GGGATGAACATAGTCTTTGTGCGGCTGAGATGCTCCCTGGAGTAAACCGGAGGACTC 366

QY 61 GCGGATGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 119
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 367 GGTGATGTTCTTGAGGACTGCCACCGCCATGCTCAATGACACAGAGTGATGAC 426
```

```
QY 120 CATTGCCCTCGCTACGACACAGTACGCTGACGCTGGGACACCTCGGTGGTCTGGGACAT 179
Db || || || || || || || || || || || || || || || || || || || || ||
427 TATAGTCCACGCTATGATCAGTACAGGATGGGTGGGATACAAATGTCCTGGCTGANIT 486
QY 180 CATGGCGGAGAAGG-----TCGGCTACTTCCACTCCATCAAGAAGGGCGTGCA 227
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
487 AAAAGTTGGAATGAAATTCARAGAATCCGCTTCTTCCACTGTATNAANAAGAAITGA 546
QY 228 CGCGGTGTGGATTGACCA-CCCGTGTCTCTGGCCCAAGGTCTGGGGCAAGACCGGCTCCA 286
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
547 CAGGGTTTTCAATTGATCATCCTTTGTTTCTTGAANAAGTGTGGGAAAACTGGANGAA 606
QY 287 AGCTGTACGGCCCCCGCTCGGGCGCTGACTA 317
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
607 TGATATNTGTCCTGTCCNGGAACGGATTA 637
```

Search completed: June 6, 2003, 22:50:16
Job time : 135.284 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:49:45 ; Search time 23.2306 Seconds
(without alignments)
2512.365 Million cell updates/sec

Title: US-09-980-771A-7

Perfect score: 2288

Sequence: 1 ALDIYVAAEVAWPSKTKGL.....AMHYGIVPVVASTGGLVDTV 438

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1257.5	55.0	606	AA25476	Granule-bound star
2	1249.5	54.6	502	AA25474	Central fragment f
3	1242	54.3	527	AA25473	Arabidopsis thalia
4	1242	54.3	563	AA25472	Arabidopsis thalia
5	1242	54.3	610	AA25471	Arabidopsis thalia
6	1242	54.3	610	AA25470	Arabidopsis thalia
7	1229	53.7	609	AA25469	Herbicidally activ
8	1185	51.8	533	AA25468	Oryza sativa starc
9	1176	51.4	637	AA25467	Zea mays waxy gene
10	870	38.0	626	AA25466	Protein encoded by
					Soluble rice starc

11	862	37.7	647	20	AA25465	Wheat starch solub
12	862	37.7	671	19	AA25464	Wheat soluble star
13	862	37.7	756	21	AA25463	Wheat soluble star
14	861	37.6	649	19	AA25462	Maize starch synth
15	859	37.5	539	19	AA25461	Zea mays starch sy
16	859	37.5	583	19	AA25460	Zea mays soluble s
17	856	37.4	792	23	AA25459	Herbicidally activ
18	850	37.2	583	19	AA25458	Herbicidally activ
19	846.5	37.0	652	23	AA25457	Herbicidally activ
20	817	35.7	799	19	AA25456	Wheat granule-boun
21	816	35.7	802	23	AA25455	Modified barley st
22	816	35.7	813	23	AA25454	Barley cultivar Mo
23	815	35.6	812	23	AA25453	Barley line MK6827
24	813	35.5	798	21	AA25452	Wheat starch synth
25	813	35.5	799	21	AA25451	Wheat starch synth
26	813	35.5	799	21	AA25450	Wheat starch synth
27	810	35.4	698	19	AA25449	Zea mays soluble s
28	810	35.4	804	19	AA25448	Maize starch solub
29	801	35.0	597	21	AA25447	Wheat starch synth
30	794.5	34.7	812	23	AA25446	Barley line 292 st
31	762.5	33.3	534	21	AA25445	Wheat starch synth
32	719	31.4	466	23	AA25444	Rice starch synth
33	696	30.4	459	17	AA25443	Soluble starch syn
34	671.5	29.3	669	19	AA25442	Maize starch solub
35	671.5	29.3	669	19	AA25441	Zea mays soluble s
36	622.5	27.2	476	23	AA25440	Streptococcus poly
37	601	26.3	483	23	AA25439	Lactococcus lactis
38	558	24.4	477	13	AA25438	glgA, Escherichia
39	558	24.4	477	19	AA25437	E. coli glycogen s
40	528	23.1	735	18	AA25436	Z. mays starch syn
41	528	23.1	677	17	AA25435	Soluble starch syn
42	528	23.1	1230	18	AA25434	Potato tuber solub
43	528	23.1	1230	21	AA25433	Potato starch synt
44	520	22.7	495	19	AA25432	Protein encoded by
45	520	22.7	1025	23	AA25431	Herbicidally activ

ALIGNMENTS

RESULT 1
AA25476
ID AA25476 standard; Protein; 606 AA.
XX
AC AA25476;
XX
DT 15-JAN-1993 (first entry)
XX
DE Granule-bound starch synthase of potato.
XX
KW GBSS gene; amylopectin; amylose production;
XX inhibition; promoter; antisense construct.
XX
OS Solanum tuberosum.
XX
PN WO9211376-A.
XX
PD 09-JUL-1992.
XX
PF 20-DEC-1991; 91WO-SE00892.
XX
PR 21-DEC-1990; 90SE-0004096.
XX
PA (AMYL-) AMYLOGENE HB.
XX
PI Hofvander P, Persson PT, Tallberg A, Wikstrom O;
XX
DR WPI; 1992-250096/30.
XX
PT N-ESDB; AAQ26404.
XX
PT Modifying potatoes to form amylopectin starch - using an
XX anti-sense construct to inhibit granule-bound starch synthase

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127482.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137582.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.

```
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 54.3%; Score 1242; DB 21; Length 527;
Best local similarity 55.3%; Pred. No. 8e-104;
Matches 250; Conservative 59; Mismatches 107; Indels 36; Gaps 9;

QY 2 LDIVAAAEVAPWSKTGGLGVTGGLPTLVKRGHRVMTIAPRYDQYADADWTSVVDI- 60
DB 1 MSVIFIGAEVGPWSKTGGLGVLGSLPPALAAARGHRVMTICPRYDQYKDAWDTCVVQIK 60
QY 61 MGEK---VRYPHSTKKGVHRWIDHPFLAKVWGTGSKLYGPRSGADYLDNHKRFALPC 117
DB 61 VGDKVENRFFPHCYKRGVDRVFDHPFLAKVVGKTGSKYKIPITGVVDYNDNQLRFLSLIC 120
QY 118 KAALEAAARVLPF-----GP-GEDCVFVANDWHSALVPVLLKDEYQPKGQFTKAKSVLA 169
DB 121 QAALEAPVNLNLSKYSFGYGEDVFEVANDWHTALLPCYLSKMSIQSGVYTMNAKVFC 180
QY 170 IHNIAFGRMWEEAFKDTKLP---QAADFKLAFSDGYAKVYTEATPMEDEKPLTGTXY 226
DB 181 IHNIAYGRAFDDYSLNLPISFKSFD---FMDGYEK-----PVKG--- 220
QY 227 KKNLWKGITAAUDLVTVTSNYATEIADRAAGGVELDTVTRAKGIEGVNMGDTEENWP 286
DB 221 KKNMKAAILFAHRLVSPYQAQELISGVDGRGVDELKYLKRMKTVSGIINGMDVOEWNP 280
QY 287 KTDKELSAFYDQNSYVAKAAKALQALQELGVPDPTAPLFAFIFGRLEEKGVDIILAAL 346
```

```
DB 281 STDKYIDIKYDITITVDKPLIKEALQAAVGLPVRDRDEVIGFIRGLEEQKGSIDLVEAI 340
QY 347 PKILATPKVQIAILGTGKAAVEKLYNAIGTKYKGRAGVVKESAPLAHMLTAGADFMVLP 406
DB 341 SKFMGL-NVQWVILGTGKKKEAQILEEKEFGKAVGKFNVPPLAHMITAGADFIIVP 399
QY 407 SRFEPCGLIQLHAMHYGIVPVVASTGGGLVDTV 438
DB 400 SRFEPCGLIQLHAMRYGTVPVIVASTGGGLVDTV 431
```

RESULT 4

```
AAG04667
ID AAG04667 standard; Protein; 563 AA.
XX AC AAG04667;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 781.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 28-APR-1999; 99US-0130891.
XX PR 30-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 04-MAY-1999; 99US-0132407.
XX PR 05-MAY-1999; 99US-0132484.
XX PR 06-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 07-MAY-1999; 99US-0132487.
XX PR 11-MAY-1999; 99US-0132863.
XX PR 14-MAY-1999; 99US-0134256.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 18-MAY-1999; 99US-0134370.
XX PR 18-MAY-1999; 99US-0134768.
XX PR 19-MAY-1999; 99US-0134941.
XX PR 20-MAY-1999; 99US-0135124.
XX PR 21-MAY-1999; 99US-0135353.
XX PR 24-MAY-1999; 99US-0135629.
XX PR 25-MAY-1999; 99US-0136021.
XX PR 27-MAY-1999; 99US-0136392.
XX PR 28-MAY-1999; 99US-0136782.
XX PR 01-JUN-1999; 99US-0137222.
XX PR 03-JUN-1999; 99US-0137528.
XX PR 04-JUN-1999; 99US-0137502.
XX PR 07-JUN-1999; 99US-0137724.
```

PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143325.
PR 14-JUL-1999; 99US-0144331.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145911.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 54.3%; Score 1242; DB 21; Length 563;
Best Local Similarity 55.3%; Pred. No. 8.8e-104;
Matches 250; Conservative 59; Mismatches 107; Indels 36; Gaps 9;

```
QY      2 LDIVMAAEVAPWSKGGGLDVTGGPLIELVKRGHRVMTIAPRYDQYADAWDTSVVVDI- 60
Db      37 MSVIFGAEEVGPWSKGGGLDVLGGPLPALAARGHRVMTICPRYDQYKDAWDTCVVQIK 96
QY      61 MGEK---VRFHSIKKGVRHVDHDFLAKVWGKTSKLYGPRSGADYLDNHRKRFALFC 117
Db      97 VGDKVENVRFHCYKRGVDFVVDHPTFLAKVWGKTSKLYGPTGYDNDNQRLRFSLLC 156
QY     118 KAATEAARVLPE-----GP-GEDCVFVANDHSAALVPVLLKDEYOPKQFTKAKSVLA 169
Db     157 QAALAPQVNLNLSKTSFGPYGEDVVFVANDHHTALLPCYLSKMSYSGRYMNAKVVC 216
QY     170 IHNIAFOGRMEAEFKDKLP---QAAFDKLAFSOGYAKVYTEATPMEDEKPLLTQKTY 226
Db     217 IHNIAFGRAFDYSLNLPISFKSFD---FMDGYEK-----PVKG--- 256
QY     227 KKNLWKGGLIADKLVTPSPNATEIAADAAGVELDVTIRAKGIGVNGMDIEFWNP 286
Db     257 RKNINWKAALIEAHRVLTVPYIAQELISGVDRGVELHKLRLMKTVSGIINGMDVQWNP 316
QY     287 KTDKFLSAPYDQNSVYAGKAAKEALQAEGLPVDPTAPLFAFTGRLEEKGVDIILAAL 346
Db     317 STDKYIDIKYDITVTDAKPLIKEALQAAVGLPVDPRDVPVIGFGRLEEKGSDILVEAI 376
QY     347 PKILATPKVQIALIGTGKAAEKLVAIGTKYKGRAGVVKFSAPLAHMLTAGADFMVLP 406
Db     377 SKFMGL-NVQMVLIGTGKKMEAQILEEKFPGKAVGVAKFNYPPLAHMTAGADFTIVP 435
QY     407 SRPEPCGLIOLHAMHYGTVPVASTGGGLVDTV 438
Db     436 SRPEPCGLIOLHAMRYGTVPVASTGGGLVDTV 467

RESULT 5
AAG04666
ID  AAG04666 standard; Protein; 610 AA.
XX
AC  AAG04666;
XX
DT  17-OCT-2000 (first entry)
XX
DE  Arabidopsis thaliana protein fragment SEQ ID NO: 780.
XX
KW  Protein identification; signal transduction pathway; metabolic pathway;
KW  hybridisation assay; genetic mapping; gene expression control; promoter;
KW  termination sequence.
XX
OS  Arabidopsis thaliana.
XX
PN  EP1033405-A2.
XX
PD  06-SEP-2000.
XX
PF  25-FEB-2000; 2000EP-0301439.
XX
PR  25-FEB-1999; 99US-0121825.
PR  05-MAR-1999; 99US-0123180.
PR  09-MAR-1999; 99US-0123548.
PR  23-MAR-1999; 99US-0125788.
PR  25-MAR-1999; 99US-0126264.
PR  29-MAR-1999; 99US-0126785.
PR  01-APR-1999; 99US-0127462.
PR  06-APR-1999; 99US-0128234.
PR  08-APR-1999; 99US-0128714.
PR  16-APR-1999; 99US-0129845.
PR  19-APR-1999; 99US-0130077.
PR  21-APR-1999; 99US-0130449.
PR  23-APR-1999; 99US-0130510.
PR  23-APR-1999; 99US-0130891.
PR  28-APR-1999; 99US-0131449.
PR  30-APR-1999; 99US-0132048.
PR  30-APR-1999; 99US-0132407.
PR  04-MAY-1999; 99US-0132484.
PR  05-MAY-1999; 99US-0132485.
PR  06-MAY-1999; 99US-0132486.
PR  06-MAY-1999; 99US-0132487.
PR  07-MAY-1999; 99US-0132863.
PR  11-MAY-1999; 99US-0134256.
PR  14-MAY-1999; 99US-0134218.
PR  14-MAY-1999; 99US-0134219.
PR  14-MAY-1999; 99US-0134221.
PR  14-MAY-1999; 99US-0134370.
PR  18-MAY-1999; 99US-0134768.
PR  19-MAY-1999; 99US-0134941.
PR  20-MAY-1999; 99US-0135124.
PR  21-MAY-1999; 99US-0135353.
PR  24-MAY-1999; 99US-0135629.
PR  25-MAY-1999; 99US-0136021.
PR  27-MAY-1999; 99US-0136392.
PR  28-MAY-1999; 99US-0136782.
PR  01-JUN-1999; 99US-0137222.
PR  03-JUN-1999; 99US-0137528.
PR  04-JUN-1999; 99US-0137502.
PR  07-JUN-1999; 99US-0137724.
PR  08-JUN-1999; 99US-0138094.
PR  10-JUN-1999; 99US-0138540.
PR  10-JUN-1999; 99US-0138847.
PR  14-JUN-1999; 99US-0139119.
PR  16-JUN-1999; 99US-0139452.
PR  16-JUN-1999; 99US-0139453.
PR  17-JUN-1999; 99US-0139492.
PR  18-JUN-1999; 99US-0139454.
PR  18-JUN-1999; 99US-0139455.
PR  18-JUN-1999; 99US-0139456.
PR  18-JUN-1999; 99US-0139457.
PR  18-JUN-1999; 99US-0139463.
PR  18-JUN-1999; 99US-0139458.
PR  18-JUN-1999; 99US-0139459.
PR  18-JUN-1999; 99US-0139460.
PR  18-JUN-1999; 99US-0139461.
PR  18-JUN-1999; 99US-0139462.
PR  18-JUN-1999; 99US-0139463.
PR  18-JUN-1999; 99US-0139750.
PR  18-JUN-1999; 99US-0139763.
PR  21-JUN-1999; 99US-0139817.
PR  22-JUN-1999; 99US-0139899.
PR  23-JUN-1999; 99US-0140353.
PR  23-JUN-1999; 99US-0140354.
PR  24-JUN-1999; 99US-0140695.
PR  28-JUN-1999; 99US-0140823.
PR  29-JUN-1999; 99US-0140991.
PR  30-JUN-1999; 99US-0141287.
PR  01-JUL-1999; 99US-0141842.
PR  01-JUL-1999; 99US-0142154.
PR  02-JUL-1999; 99US-0142055.
PR  06-JUL-1999; 99US-0142390.
PR  08-JUL-1999; 99US-0142803.
PR  09-JUL-1999; 99US-0142920.
PR  12-JUL-1999; 99US-0142977.
PR  13-JUL-1999; 99US-0143542.
PR  14-JUL-1999; 99US-0143624.
PR  15-JUL-1999; 99US-0144005.
PR  16-JUL-1999; 99US-0144085.
PR  16-JUL-1999; 99US-0144086.
PR  19-JUL-1999; 99US-0144325.
PR  19-JUL-1999; 99US-0144331.
PR  19-JUL-1999; 99US-0144332.
PR  19-JUL-1999; 99US-0144333.
PR  19-JUL-1999; 99US-0144334.
PR  19-JUL-1999; 99US-0144335.
PR  20-JUL-1999; 99US-0144352.
PR  20-JUL-1999; 99US-0144632.
PR  20-JUL-1999; 99US-0144884.
PR  21-JUL-1999; 99US-0144814.
PR  21-JUL-1999; 99US-0145086.
PR  21-JUL-1999; 99US-0145088.
PR  22-JUL-1999; 99US-0145085.
```


PR 12-SEP-1997; 97AU-0009108.
 XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
 PA (GOOD-) GOODMAN FIELDER LTD.
 PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
 XX (AUSU) UNIV AUSTRALIAN NAT.
 PI Li Z, Morell M, Rahman S;
 XX WPI; 1999-229525/19.
 DR N-PSDB; AAX34651, AAX34652.
 XX
 PT New isolated cereal plant enzyme genes used for, e.g. expression of
 PT antisenase sequences of granule bound synthase
 XX
 PS Claim 13; Page 95-97; 17lpp; English.
 XX
 CC The invention relates to a novel enzyme of starch biosynthetic pathway
 CC in a cereal plant, where the enzyme is selected from starch branching
 CC enzyme (SBE) I, SBE II, starch soluble synthase (SSS) I, and debranching
 CC enzyme (DBE), with the proviso that the enzyme is not SSS I of rice, or
 CC SBE I of rice or maize. The methods and products can be used for
 CC targeting expression specifically to the endosperm of the seeds of cereal
 CC plants such as wheat or barley. They can be used for the expression of
 CC e.g. antisenase sequences of granule-bound synthase (GBSS), SBE II, low
 CC mol. wt. glutenin, grain softness protein I, bacterial isoamylase,
 CC bacterial glycogen synthase, and wheat high mol. wt. glutenin Bx17. They
 CC can be used for modifying the characteristics of starch produced by a
 CC plant. The present sequence represents the wheat SSS I protein sequence.
 XX
 SQ Sequence 647 AA;
 Query Match 37.7%; Score 862; DB 20; Length 647;
 Best Local Similarity 43.1%; Pred. No. 3.1e-69;
 Matches 194; Conservative 70; Mismatches 148; Indels 38; Gaps 12;
 QY 4 IVVAAEAVAPSKTGGLDVTGGPIELVKGHRVMTAPRY-----DQYADAWDTSVV 57
 DB 141 IVFTGGAAPYAKSGGLDVCGLPIALAAARGHRVVMVPRYLNSSDKNTAKALYTKGH 200
 QY 58 VDIM----GEKRVFHSIKKGVRWIDHPWFLAKVWGKTGSKLYGPRSGADYLDNKKRF 113
 DB 201 IKIPCFGSHVETFFHEYRDNVDWVFDHP-----SYHRPGS-LYGDNFGA-FGDNQPRY 253
 QY 114 ALFCKAAEAARVLFPFGP---GECVFNVDHWSALVPLVKDYPQKQFTKAKSLVLA 170
 DB 254 TLLCYAAACEAPLILELGGYIYGQNCMFVNDHWSALVPLVLAARYPYGVYRDSRSTLVI 313
 QY 171 HNIAPQGRMWEAEAFKDKLPQAAFDKLAFLSDGYAKVYTEATPMEDEKPLTGTYYKIN 230
 DB 314 HNLHQGLEPASTYFDLGLPEWYGALEW-----VPEWARRIALDKG-----EAVN 360
 QY 231 WLKGGIIAADKLVTVSNNYAFETIAADAGGVYELDTVI--RAKGIEGIVNGMDIEWNPKT 288
 DB 361 FLKGVVTVADRIVVSQCYSEVTT-ABGGQGLNELLSSRKSVLNGIVNGIDINDWNPTT 419
 QY 289 DKFLSAPYQNSVYAGRAAKAEALQAEGLPVDPTAPLFAFIEGLEQKGVDDIILALPK 348
 DB 420 DKCLPHHYSVDDL-SGAKCKAEALQELGLPVEDVPLIGFIRGLDYQKGLDLKMAIPE 478
 QY 349 ILAIPKQVQIALIGTGKAAEKLVAIGTKYGRKGVVVKFSAPLAHMLTAGADEMLVPSR 408
 DB 479 LM-REDVQFVMLGSGDPTFEQWMSRSTSSYKDKRGVGVFSVPVSHRITAGCDILLMPSR 537
 QY 409 FEPGLIQLHAMHYGTVPVWASTGGLVDVT 438
 DB 538 FEPGLNOLYAMQYGTVPVWHTGGLRTV 567
 RESULT 12
 AAW23937
 ID AAW23937 standard; Protein; 671 AA.
 XX

AC AAW23937;
 XX 21-MAY-1998 (first entry)
 XX Wheat soluble starch synthase partial sequence.
 DE Starch synthase; wheat; transgenic plant.
 XX Starch synthase; wheat; transgenic plant.
 XX Triticum aestivum L. cv. Florida.
 OS WO9745545-A1.
 PN 04-DEC-1997.
 XX 28-MAY-1997; 97WO-EP02793.
 XX 11-SEP-1996; 96DE-1036917.
 PR 29-MAY-1996; 96DE-1021588.
 XX (AGRE) HOECHST-SCHERING AGREVO GMBH.
 PA Block M, Loerz H, Luetticke S, Froberg C, Kossmann J;
 PI Walter L;
 XX WPI; 1998-032652/03.
 DR N-PSDB; AAV01527.
 XX
 PT Nucleic acid encoding starch synthase enzymes from wheat - for
 PT transgenic plants that produce modified forms of starch, useful e.g.
 PT in foods, or for production of packaging materials and disposable
 PT goods
 XX
 PS Claim 1; Page 47-50; 7lpp; English.
 XX
 CC This amino acid sequence comprises a near full-length sequence for
 CC a soluble starch synthase of summer wheat (cv. Florida). It was
 CC deduced from a cDNA clone (AAV01527) isolated from a 21-day caryopsis
 CC cDNA library. A granule-bound starch synthase (see AAW23938) has
 CC also been identified. Isolated nucleic acids encoding these
 CC enzymes can be inserted into vectors for production of transgenic
 CC plants, particularly starch-producing plants, specifically wheat.
 CC Use of the isolated nucleic acids, or of antisense sequences, allows
 CC starch metabolism to be regulated in transgenic plants.
 CC Overexpression may result in improved crop yield, while modification
 CC of starch in planta may eliminate the need for subsequent
 CC chemical/physical modification. Plants with altered levels of the
 CC various isoforms of starch synthase will produce starch of different
 CC chain length, amylose/amylopectin ratio, degree of branching,
 CC phosphate content, gelatinisation behaviour, granule size and shape,
 CC viscosity etc. The starch produced by such plants is useful
 CC particularly in foods (especially bakery goods or pasta) or to
 CC produce packaging materials or disposable goods, as well as in any
 CC other known use of starch.
 XX
 SQ Sequence 671 AA;
 Query Match 37.7%; Score 862; DB 19; Length 671;
 Best Local Similarity 43.1%; Pred. No. 3.2e-69;
 Matches 194; Conservative 71; Mismatches 147; Indels 38; Gaps 12;
 QY 4 IVVAAEAVAPSKTGGLDVTGGPIELVKGHRVMTAPRY-----DQYADAWDTSVV 57
 DB 56 IVFTGGAAPYAKSGGLDVCGLPIALAAARGHRVVMVPRYLNSSDKNTAKALYTKAH 115
 QY 58 VDIM----GEKRVFHSIKKGVRWIDHPWFLAKVWGKTGSKLYGPRSGADYLDNKKRF 113
 DB 116 IKIPCFGSHVETFFHEYRDNVDWVFDHP-----SYHRPGS-LYGDNFGA-FGDNQPRY 168
 QY 114 ALFCKAAEAARVLFPFGP---GECVFNVDHWSALVPLVKDYPQKQFTKAKSLVLA 170
 DB 169 TLLCYAAACEAPLILELGGYIYGQNCMFVNDHWSALVPLVLAARYPYGVYRDSRSTLVI 228
 QY 171 HNIAPQGRMWEAEAFKDKLPQAAFDKLAFLSDGYAKVYTEATPMEDEKPLTGTYYKIN 230

Db 229 HNLHQGVPEASTYDGLPPEWYGALEW-----VFPEWARRHALDKG-----EAVN 275
 QY 231 WLKGGITIAADKLVTVSPNYATEIAAAGGVVELDTVI--RAKGTEGIVNGMDIEWNPKT 288
 Db 276 FLKGVVTVADRIIVTSQGYSEWVIT-AGGQGLNELSSRSKSVLNGVINGIDINDWNPPT 334
 QY 289 DKFLSAPYDQNSVYAGKAAKEALQAEGLPVDPTAPLFAFIGNLEBQKGVDIILAAALPK 348
 Db 335 DKCLPHHYSVDDL-SGKAKCAELQKELGLPVREDVPLIGFIGNLDYQKGLDLKMAIPE 393
 QY 349 ILATPKVQIAILGTGKAAYEKLVNAIGTKYGRAGKGVVVKFSAPLAHMLTAGADMVPSR 408
 Db 394 LM-REDVQFVWLGSGDPIFEQWMRSTESSYKDKFRGWGVFSPVSHRITAGCDILLMPSR 452
 QY 409 FEPGLQLHAMHYGTVPVASTGGLVDTV 438
 Db 453 FEPGLNQLYAMQYGTVPVHGTGLRDTV 482

RESULT 13

AAV50818
 ID AAV50818 standard; Protein; 756 AA.

XX AC AAV50818;
 XX 18-FEB-2000 (first entry)

DE DT Wheat soluble starch synthase protein.
 DE DE Soluble; starch synthase; wheat; transgenic plant; starch production;
 KW food; baking; pastry; packaging material; glucose; glucan; paper; pulp;
 KW adhesive; textile; building material; soil stabilizer; wetting agent;
 KW fertilizer; plant-protection; cosmetic; flocculant.

XX OS Triticum aestivum.

XX FN DE19820607-A1.

XX PD 11-NOV-1999.

XX PF 08-MAY-1998; 98DE-1020607.

XX PR 08-MAY-1998; 98DE-1020607.

XX PA (AGRE) HOECHST-SCHERING AGREVO GMBH.

XX PI Loerz H, Luetticke S, Block M;

XX WPI; 2000-024508/03.

XX DR N-PSDB; AA224487.

XX PT New enzyme with starch synthase activity, useful for producing starch
 for foods and packaging materials -
 XX Claim 1a; Page 19-21; 24pp; German.

XX This invention describes a novel protein (I) with the activity of wheat
 starch synthase. Transgenic plants, specifically wheat, that contain (I)
 are used for production of starch, used particularly in foods,
 CC particularly baked and pastry goods and for making packaging materials or
 CC disposable items. Starch may also be used as starting materials for
 CC glucose or glucan components (e.g. for fermentation or further chemical
 CC conversion); in paper and pulp production, as adhesives, in textiles,
 CC in preparation of gypsum-based building materials, as soil stabilizer,
 CC as wetting agent etc. in fertilizer and plant-protection compositions,
 CC as binder (in pharmaceuticals, cosmetics, coal briquetting and casting
 CC sand), as flocculant in soil or coal slurries, as rubber and leather
 CC additives, and for production of synthetic polymers, e.g. polyurethane
 CC films. Transgenic plants with increased/decreased production of (I)
 CC produce starches with altered physical and/or chemical properties such as
 CC amylose/amylopectin ratios, degree of branching, mean chain length,
 CC phosphate content, gelatinization properties, gel- or film-forming

CC properties, or starch grain size or structure. This sequence represents
 CC the soluble starch synthase isolated from wheat (Triticum aestivum L. cv.
 CC Florida).
 XX Sequence 756 AA;
 SQ Query Match 37.7%; Score 862; DB 21; Length 756;
 Best Local Similarity 43.1%; Pred. No. 3.8e-69;
 Matches 194; Conservative 71; Mismatches 147; Indels 38; Gaps 12;
 QY 4 IVMTAAEVAPSKTGGGLDVTGGLPIELVKGHRVMTIAPRY-----DOYADAWDSVV 57
 Db 141 IVFTVGEAAPYAGSGGLDVCGLPIALAAARGHRVWVMPYRLNGSSDKNTAKALYAKH 200
 QY 58 VDIM---GEKRVYFHSIKKGVRWIDHPWFLAKVNGKTGSKLYGPRSGADYLDNHKRF 113
 Db 201 IKIPFCGSGHEVTFEYRONVWVVDHP-----SYHRPGS-LYGDNFGA-FGDNQFRY 253
 QY 114 ALFCKAAATEARVLPFGP---GECVVFVANDWHSALVPVLLKDEYQPKGQTKAKSVLAI 170
 Db 254 TLLCYAAACEAPLILELGGYIYGQNCMFVYNDWHASLVPVLLAAKRYPYGVYRDSRSTLVI 313
 QY 171 HNIAFQGRMEEAFKDKLPQAAFDKLAFSQYAKVYTEATPMEEDKPPLTGKTYKKIN 230
 Db 314 HNLHQGVPEASTYDGLPPEWYGALEW-----VFPEWARRHALDKG-----EAVN 360
 QY 231 WLKGGITIAADKLVTVSPNYATEIAAAGGVVELDTVI--RAKGTEGIVNGMDIEWNPKT 288
 Db 361 FLKGVVTVADRIIVTSQGYSEWVIT-AGGQGLNELSSRSKSVLNGVINGIDINDWNPPT 419
 QY 289 DKFLSAPYDQNSVYAGKAAKEALQAEGLPVDPTAPLFAFIGNLEBQKGVDIILAAALPK 348
 Db 420 DKCLPHHYSVDDL-SGKAKCAELQKELGLPVREDVPLIGFIGNLDYQKGLDLKMAIPE 478
 QY 349 ILATPKVQIAILGTGKAAYEKLVNAIGTKYGRAGKGVVVKFSAPLAHMLTAGADMVPSR 408
 Db 479 LM-REDVQFVWLGSGDPIFEQWMRSTESSYKDKFRGWGVFSPVSHRITAGCDILLMPSR 537
 QY 409 FEPGLQLHAMHYGTVPVASTGGLVDTV 438
 Db 538 FEPGLNQLYAMQYGTVPVHGTGLRDTV 567

RESULT 14

AAW38218
 ID AAW38218 standard; Protein; 649 AA.

XX AC AAW38218;

XX DT 22-MAY-1998 (first entry)

XX DE Maize starch synthase type I.

XX KW Maize; starch synthase type I; starch.

XX OS Zea mays.

XX FN DE19619918-A1.

XX PD 20-NOV-1997.

XX PF 17-MAY-1996; 96DE-1019918.

XX PR 17-MAY-1996; 96DE-1019918.

XX PA (PLAN-) PLANITEC BIOTECHNOLOGIE GMBH.

XX PI Froberg C, Kossmann J;

XX DR WPI; 1998-000821/01.

XX DR N-PSDB; AAT95785.

XX PT DNA encoding maize starch synthase type I protein - for producing

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 14:47:01 ; Search time 11.3212 Seconds
(without alignments)
3719.278 Million cell updates/sec

Title: US-09-980-771A-7
Perfect score: 2288
Sequence: 1 ALDIVMVAEVPWSKTGGL.....AMHYGTVPVASTGGLVDTV 438

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1301	56.9	608	2 S43341	starch synthase (E
2	1275	55.7	607	1 YUPOY	starch synthase (E
3	1258	55.0	603	2 S61504	glycogen(starch) s
4	1246	54.5	608	2 T10906	starch synthase (E
5	1242	54.3	610	2 F86453	granule-bound star
6	1232.5	53.9	603	1 YUBHY	glycogen(starch) s
7	1230	53.8	609	2 JQ0703	glycogen(starch) s
8	1229	53.7	609	2 S11481	glycogen(starch) s
9	1204	52.6	615	1 YUWY	glycogen(starch) s
10	1187	51.9	608	2 T14731	glycogen(starch) s
11	1185	51.8	605	1 S07314	glycogen(starch) s
12	870.5	38.0	641	2 T07668	starch synthase (E
13	870	38.0	626	2 JQ2322	starch synthase (E
14	859	37.5	622	2 T01414	starch synthase (E
15	850	37.2	610	2 T06280	probable starch sy
16	833	36.4	752	2 S61505	glycogen(starch) s
17	822	35.9	732	2 T01208	starch synthase (E
18	815	35.6	491	2 T06798	probable starch sy
19	810	35.4	698	2 T01209	starch synthase (E
20	805.5	35.2	788	2 T07667	glycogen(starch) s
21	660.5	28.9	477	2 B95130	glycogen synthase
22	659.5	28.8	477	2 H98000	starch synthase (E
23	628	27.4	484	2 S40051	starch synthase (E
24	623.5	27.3	480	2 H98228	glycogen synthase
25	623.5	27.3	480	2 AF3057	glycogen synthase
26	611	26.7	486	2 H72321	glycogen synthase
27	601	26.3	477	2 A97176	glycogen synthase
28	601	26.3	483	2 C86712	starch synthase (E
29	589	25.7	484	2 F82165	glycogen synthase

30	581	25.4	476	2 AE0479	starch synthase (E
31	578.5	25.3	476	2 E83785	starch (bacterial
32	570.5	24.9	486	2 C96018	probable starch sy
33	566	24.7	477	2 AI0995	starch synthase (E
34	565	24.7	477	1 SYEGGL	starch synthase (E
35	565	24.7	477	2 B98163	glycogen synthase
36	565	24.7	477	2 C86009	glycogen synthase
37	541	23.6	530	2 AH3194	soluble starch syn
38	528	23.1	1230	2 T07663	glycogen synthase
39	524	22.9	476	2 C64119	starch synthase (E
40	522.5	22.8	463	2 C70363	glycogen synthase
41	520.5	22.7	472	2 AI2040	glycogen synthase
42	520	22.7	1025	2 H86250	hypothetical prote
43	506.5	22.1	477	2 S76496	hypothetical prote
44	502	21.9	492	2 AG1810	glycogen (starch)
45	500.5	21.9	1071	2 T04926	starch synthase ho

ALIGNMENTS

RESULT 1

S43341

starch synthase (EC 2.4.1.21) precursor - cassava

N:Alternate names: bacterial-glycogen synthase; glycogen synthase; starch synthase
C:Species: Manihot esculenta (cassava)

C>Date: 07-Sep-1994 #sequence_revision 03-Nov-1995 #text_change 16-Aug-2002

C:Accession: S43341

R:Salehuzzaman, S.N.I.M.; Jacobsen, E.; Visser, R.G.F.

Plant Mol. Biol. 23, 947-962, 1993

A:Title: Isolation and characterization of a cDNA encoding granule-bound starch synth

A:Reference number: S43341; MUID:94083565; PMID:8260633

A:Accession: S43341

A:Molecule type: mRNA

A:Residues: 1-608 <SAL>

A:Cross-references: EMBL:X74160; NID:g437041; PIDN:CAA52273.1; PID:g437042

C:Genetics:

A:Gene: GBSS; waxy

A:Genome: nuclear

C:Function:

A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producin

C:Superfamily: starch synthase

C:Keywords: amyloplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltra

F:1-78/Domain: transit peptide (amyloplast) #status predicted <TNP>

F:79-608/Product: ADPglucose-starch glucosyltransferase #status predicted <MAT>

Query Match 56.9%; Score 1301; DB 2; Length 608;

Best Local Similarity 57.5%; Pred.No.3.7e-85;

Matches 258; Conservative 51; Mismatches 110; Indels 30; Gaps 7;

QY	2	LDIVMVAEVPWSKTGGLDVTGGTGLIELVKKGRHVRMTIAPRYDQYADAWDTSVVVDI- 60	
DB	82	MNLIFVGAEGVGPWSKTGGLDVLGGLPPAMAARGHVRMTIVSPRYDQYKDAWDTSVSVEIK 141	
QY	61	MG---EKVYFHSIKKGVHVRWIDHPWFLAKVWGTGSKLYGPRGADYLDNKKRALFC 117	
DB	142	IGDRIETVRFHSHYKRGVDRVDFVDFHMFLEKFWGKTGSKIYGPAGLDYQDNQLRFLSLC 201	
QY	118	KAALAEARVLPF-----GP-GEQVFNVDHWSALVPLVKDEYQPKQFTKAKSVLA 169	
DB	202	LAALAPRVNLNLSKNSGPGYEEVAFANDHALLPCYLKAIQPMGIYKHAKVAPC 261	
QY	170	IHTIAFGRMWEAFKDKLFOAADFCLAFSDGYAKVYTEATPMEEDEKPPLTGTYKKI 229	
DB	262	LHNIAVQGREAFSDPRLNLPDKFKSSDFDIDGIEK-----PVKG---RKI 304	
QY	230	NWLKGGITAAKLVTVSPNYATEIAADAAGGVVELDTVIRAKGIEGVNGMDIEWNPKTD 289	
DB	305	NMKAGILESRLVTVSPYQAQEVLSVGVGVVDFNFKTKGTGIAGIINGMDVQWNPVTD 364	
QY	290	KFLSAPYQNSVYAGKAAKALQALQELGIPVDPTAPLAFVIGRLEEQGVDIILALPKI 349	
DB	365	KYIDHYDATTVMDAKPLKKEALQAEVGLPVDNRNPLIGFVIGRLEEQGSDFVAAISQ- 423	

Db 417 AK-FADENVQIVVLGTGKKIMEKQIEVLEKEYPGKAIGITKFNPSLAHMLIAGADFIIVP 475
 QY 407 SRFEPCGLIOLHAMHYGVVAVSTGGLVDV 438
 Db 476 SRFEPCGLVQLHAMPIGIVPIVSSGTGLVDV 507

RESULT 4

T10906
 starch synthase (EC 2.4.1.21) - sweet potato
 N:Alternate names: starch synthase
 C:Species: Ipomoea batatas (sweet potato)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Aug-2002
 C:Accession: T10906

R:Wang, S.J.; Yeh, K.W.; Tsai, C.Y.
 submitted to the EMBL Data Library, December 1995
 A:Reference number: Z17212
 A:Accession: T10906
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-608 <W>
 A:Cross-references: EMBL:U44126; NID:g1172158; PID:g1172159
 A:Experimental source: cv. Tainong; tuberous root
 C:Genetics: SS67
 A:Gene: SS67
 C:Function:

A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producing a
 A:Pathway: starch biosynthesis
 C:Superfamily: starch synthase
 C:Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 54.5%; Score 1246; DB 2; Length 608;
 Best Local Similarity 56.3%; Pred. No. 3.2e-81;
 Matches 253; Conservative 48; Mismatches 118; Indels 30; Gaps 7;

QY 2 LDIVMAAEVAPWSKTGGLGVDVTGGLPIELVKKRHRVMTIAPRYDOYADAWDTSVVVDI- 60
 Db 82 MNLVFGCEGPGWCKTGGLGVDVGLPLPALAARGHRVMTICPRYDOYKDAWETCVVVPQ 141
 QY 61 MG---EKVYFHSIKKGVRHWIDHPWFLAKVWGKTGSKLYGPRGADYLDNHRKRALFC 117
 Db 142 VGDRIEPRVRFHYSKRGVDRVDFVHPMFLKVGKTSMLYGPAGKDYKDNQLRFLSLC 201
 QY 118 KAAIEAARVLPF-----GP-GEDEVFANDHWSALVPVLLKDEYQPKGQFTKAKSVLA 169
 Db 202 QAALAEAPVNLNLSKYSFGPYGEDVVFVANDHWHALLPCYIKWYQSRGIYMNKAVFC 261
 QY 170 IHNIAFGRMWEEAFKDTKLPAFAFSDGAKVYVTEATPMEDEKPPITGKTYKKI 229
 Db 262 IHNIAFGREFAFSDFLNLNLPDEYKGSFDFIDGYDK-----PVKG---RKI 304
 QY 230 NMLKGGIITADKLVTSPNYATEAADAAGVGLDVTIRAKGIEGIVNGMDEEWNPKTD 289
 Db 305 NWMKAGIREADRVETVSPNYAKELVSCVSKGVLDNHIRDCGTCGNGMDTQEWNPATD 364
 QY 290 KFLSAPYDONSVAAGKAAKALQAEGLPVDPTAPFAFICGRLEOKGVDIILALPKI 349
 Db 365 KYLAVKDYDITVMQAKPLKALQAVGLPVDRIPLIGFICGRLEOKGSDIILAAISKF 424
 QY 350 LATPKVQITAILGTGKAAEYKLVNAIGTKYGRKGVVVKFSAPLAHMLTAGADFMVPSRF 409
 Db 425 ISM-DVQIILGTGKKFEQIQELEVMPDKARGVAKFNVPYLAHMITAGADFMVPSRF 483
 QY 410 EPCGLIOLHAMHYGVVAVSTGGLVDV 438
 Db 484 EPCGLIOLHAMRYGTPCICASTGGLVDV 512

RESULT 5

F86453
 granule-bound starch synthase [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: F86453
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alor
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Marzie
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tall
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712

C:Accession: F86453
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-610 <STO>
 A:Cross-references: GB:AE005172; NID:96910568; PID:AAF31273.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1
 C:Superfamily: starch synthase

Query Match 54.3%; Score 1242; DB 2; Length 610;
 Best Local Similarity 55.3%; Pred. No. 6.1e-81;
 Matches 250; Conservative 59; Mismatches 107; Indels 36; Gaps 9;

QY 2 LDIVMAAEVAPWSKTGGLGVDVTGGLPIELVKKRHRVMTIAPRYDOYADAWDTSVVVDI- 60
 Db 84 MSVIEIGAEGVPSKTGGLGVDVGLPLPALAARGHRVMTICPRYDOYKDAWETCVVQIK 143
 QY 61 MCEK---VRYFHSIKKGVRHWIDHPWFLAKVWGKTGSKLYGPRGADYLDNHRKRALFC 117
 Db 144 VGDKVENVRFFHCYKRGVDRVDFVHPFLAKVWGKTGSKYIGFITGVNDNQLRFLSLC 203
 QY 118 KAAIEAARVLPF-----GP-GEDEVFANDHWSALVPVLLKDEYQPKGQFTKAKSVLA 169
 Db 204 QAALAEAPVNLNLSKYSFGPYGEDVVFVANDHWHALLPCYIKWYQSRGIYMNKAVFC 263
 QY 170 IHNIAFGRMWEEAFKDTKLPAFAFSDGAKVYVTEATPMEDEKPPITGKTYKKI 226
 Db 264 IHNIAFGREFAFDDYSLNLNLPISFKSFD---FMDGYEK-----PVKG--- 303
 QY 227 KKNLKGIIAADKLVTSPNYATEAADAAGVGLDVTIRAKGIEGIVNGMDEEWNPK 286
 Db 304 KKNMKAALIEAHRVLTVPSPYTAQELTSGVDRGVELHKLKMKTVSGIINGMDYQEWNP 363
 QY 287 KTDKFLSAPYDONSVAAGKAAKALQAEGLPVDPTAPFAFICGRLEOKGVDIILAL 346
 Db 364 STDKYIDIKYDITVTDAKPLKALQAVGLPVDRIPLIGFICGRLEOKGSDIILVEAI 423
 QY 347 PKLATPKVQITAILGTGKAAEYKLVNAIGTKYGRKGVVVKFSAPLAHMLTAGADFMVLP 406
 Db 424 SKFMGL-NVQMVLGTGKKMEQAQILEEEKFPKAVGVAKFNVPYLAHMITAGADFIIVP 482
 QY 407 SRFEPCGLIOLHAMHYGVVAVSTGGLVDV 438
 Db 483 SRFEPCGLIOLHAMRYGTPVIVASTGGLVDV 514

RESULT 6

YU8HY
 glycogen(starch) synthase (EC 2.4.1.11) precursor - barley
 N:Alternate names: starch synthase
 C:Species: Hordeum vulgare (barley)
 C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 02-Aug-2002
 C:Accession: S01727; S01728
 R:Roeder, W.; Becker, D.; Salamini, F.
 Nucleic Acids Res. 16, 7185-7186, 1988
 A:Title: Structural analysis of the waxy locus from Hordeum vulgare.
 A:Reference number: S01727; MUID:88303345; PMID:2970062
 A:Accession: S01727
 A:Molecule type: DNA
 A:Residues: 1-603 <ROI>
 A:Cross-references: EMBL:X07931; NID:q19126; PIDN:CAA30755.1; PID:g295809

A:Residues: 1-609 <HIR>
 A:Cross-references: EMBL:X58228; NID:g20400; PIDN:CAA1186.1; PID:g20401
 A:Accession: PC2190
 A:Molecule type: protein
 A:Residues: 78-113 <H12>
 A:Experimental source: leaf, cDNA pOSLHC2120
 C:Comment: This protein is involved in amylose synthesis in the rice endosperm.
 C:Genetics:
 A:Gene: waxy
 A:Introns: 113/3; 140/3; 173/3; 203/3; 225/1; 258/3; 295/2; 376/3; 435/3; 499/3; 528/3;
 C:Function:
 A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producing
 A:Pathway: starch biosynthesis
 C:Superfamily: starch synthase
 C:Keywords: amyloplast; chloroplast; glycogen/starch biosynthesis; glycosyltransferase;
 F:1-77/Domain: transit peptide (amyloplast) #status predicted <INP>
 F:78-609/Product: glycogen (starch) synthase #status experimental <MAT>

Query Match 53.7%; Score 1229; DB 2; Length 609;
 Best Local Similarity 53.9%; Pred. No. 5,2e-80;
 Matches 242; Conservative 61; Mismatches 116; Indels 30; Gaps 6;
 QY 2 LDIVVAAEVPWPKTGLGSDVTGGLPIELVKRHRVMTIAPRYDQYADADMTSVVVDI- 60
 DB 83 MNVVFVGAEMAPWPKTGLGSDVTGGLPAAANGHRVWVISPRIYDQYKDAWDTSVVAEIK 142
 QY 61 ---MGEKRYRPHSIKKGVRHWIDHPFLAKVWGTGSKLYGPRSGADYLDNKKRFALEFC 117
 DB 143 VADRYERVRFHCYKRGVDRVFDHPSELEKVGWTKGIYGPDTGVYKDNQMRFSLLC 202
 QY 118 KAAIEAARVL-----PFGP-----GEDCVFVANDHSAIPVLLKDEYQPKGFTAKSVLA 169
 DB 203 QAALAEAPRLNNPFGKTYGDEGVFVNCNDHWHTGLLACVLYKSNYQSN 262
 QY 170 IHNIAFOGRMWEAFKDTKLPOAAFDKLAISDGYAKVYTEATPMEDEKPPITGKTYKKI 229
 DB 263 IHNISYGRFAEDYPELNLSERFSRSPDFIDGY-----DTPVEG-----RKI 305
 QY 230 NWLKGIIAADKLVTSVSNYATEATAADAGGVDELDTVIRAKGIEGIVNGMDEWNPRTD 289
 DB 306 NMMKAGILEADRVLTVSPYAEELISGIFARGCELDNIMRLGITGIVNGMDVSEWDPSKD 365
 QY 290 KFLSAPYDQNSVYAGKAAKALQALQELGLPYDTPAPFAFVIGRLEEOKGVDIILAALEPKI 349
 DB 366 KYITAKYDATTAIAKALNKALQALQELGLPYDTPAPFAFVIGRLEEOKGVDPVMAAIPEL 425
 QY 350 LATPKVQIAIIGTGAAYEKILVNAIGTKYKGRAGVVKFSAPLAHMLTAGADEMLVPSRF 409
 DB 426 M-QEDVQVILVLTGKTKKFEKLLKMEKYPCKVRAVVKFNAPLAHMLMAGADVLAVPSRF 484
 QY 410 EPGGLIOLHAMHYGTVPVWASTGGLVDTV 438
 DB 485 EPGGLIOLQGMRYGTGTPCACASTGGLVDTV 513

RESULT 9
 YUWVY
 glycogen(starch) synthase (EC 2.4.1.11) precursor - wheat
 N:Alternate names: starch synthase
 C:Species: Triticum aestivum (common wheat)
 C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 02-Aug-2002
 C:Accession: S33636
 R:Clark, J.R.; Robertson, M.; Ainsworth, C.C.
 Plant Mol. Biol. 16, 1099-1101, 1991
 A:Title: Nucleotide sequence of a wheat (Triticum aestivum L.) cDNA clone encoding the
 A:Reference number: S16261; MUID:913222506; PMID:1863765
 A:Accession: S16261
 A:Molecule type: mRNA
 A:Residues: 1-615 <CLA>
 A:Cross-references: EMBL:X57233; NID:g21901; PIDN:CAA0509.1; PID:g21902
 R:Ainsworth, C.; Clark, J.; Baisdon, J.
 Plant Mol. Biol. 22, 67-82, 1993
 A:Title: Expression, organisation and structure of the genes encoding the waxy protein

A:Reference number: S33636; MUID:93271462; PMID:8499619
 A:Accession: S33636
 A:Molecule type: protein
 A:Residues: 71-75, 'X', 77-78 <AIN>
 C:Genetics:
 A:Gene: waxy
 C:Function:
 A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose produc
 A:Pathway: starch biosynthesis
 C:Superfamily: starch synthase
 C:Keywords: amyloplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltra
 F:1-70/Domain: transit peptide (amyloplast) #status predicted <INP>
 F:71-615/Product: UDPglucose-starch glycosyltransferase #status experimental <MAT>
 Query Match 52.6%; Score 1204; DB 1; Length 615;
 Best Local Similarity 51.7%; Pred. No. 3,2e-78;
 Matches 238; Conservative 60; Mismatches 122; Indels 40; Gaps 5;
 QY 2 LDIVVAAEVPWPKTGLGSDVTGGLPIELVKRHRVMTIAPRYDQYADADMTSVVVDI- 60
 DB 77 MNVVFVGAEMAPWPKTGLGSDVTGGLPAAANGHRVWVISPRIYDQYKDAWDTSVIASEIK 136
 QY 61 ---MGEKRYRPHSIKKGVRHWIDHPFLAKVWGTGSKLYGPRSGADYLDNKKRFALEFC 117
 DB 137 VADRYERVRFHCYKRGVDRVFDHPSELEKVGWTKGIYGPDTGVYKDNQMRFSLLC 196
 QY 118 KAAIEAARVLPP-----GP-----GEDCVFVANDHSAIPVLLKDEYQPK 158
 DB 197 QAALAEAPRLNNPFGSPVAMLCRAVPRAGEDVVFVNCNDHWHTGLLACVLYKSNYQSN 256
 QY 159 GQFTAKSVLAHNIAFOGRMWEAFKDTKLPOAAFDKLAISDGYAKVYTEATPMEDEK 218
 DB 257 GIYRTAKVAFCHNLSYGRFSFDDFAQLNLPDREKSSPFDIDGYDK----- 303
 QY 219 PELTGTCKYKINWLGKGIITADKLVTSVSNYATEATAADAGGVDELDTVIRAKGIEGIVNG 278
 DB 304 -PVEG---RKINWKGILQADKLVTSVSPYAEELISGEARGCELDNIMRLGITGIVNG 359
 QY 279 MDIEWNPRTDKLSAPYDQNSVYAGKAAKALQALQELGLPYDTPAPFAFVIGRLEEOKG 338
 DB 360 MDVSEMDPKDKFLVNDVTTALGKALNKALQALQELGLPYDTPAPFAFVIGRLEEOKG 419
 QY 339 VDIILAALEPKIATPKVQIAIIGTGAAYEKILVNAIGTKYKGRAGVVKFSAPLAHMLTA 398
 DB 420 PDVMAIATPEIVKEDVQVILVLTGKTKKFEKLLKMEKYPCKVRAVVKFNAPLAHOMMA 479
 QY 399 GADFMVPSRPEPCGLIOLHAMHYGTVPVWASTGGLVDTV 438
 DB 480 GADVLAVTSRPEPCGLIOLQGMRYGTGTPCACASTGGLVDTV 519

RESULT 10
 TI4731
 glycogen(starch) synthase (EC 2.4.1.11) precursor, granule-bound - sorghum
 N:Alternate names: starch synthase, granule-bound
 C:Species: Sorghum bicolor (sorghum)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2002
 C:Accession: T14731
 R:Hsing, Y.C.
 submitted to the EMBL Data Library, April 1995
 A:Reference number: Z18173
 A:Accession: T14731
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-608 <HST>
 A:Cross-references: EMBL:U23945; NID:gl255713; PID:gl255714
 A:Experimental source: strain 12311; young seed
 C:Genetics:
 A:Genome: nuclear
 A:Note: Wx
 C:Function:
 A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose produc
 A:Pathway: starch biosynthesis

C;Superfamily: starch synthase
C;Keywords: amyloplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase
F;1-77/Domain: transit peptide (amyloplast) #status predicted <TMP>
F;76-608/Product: UDPglucose-starch glucosyltransferase #status predicted <MAT>

Query Match 51.9%; Score 1187; DB 2; Length 608;
Best Local Similarity 52.6%; Pred. No. 5.2e-77;
Matches 236; Conservative 63; Mismatches 120; Indels 30; Gaps 8;

```
QY 2 LDIVMAAEVAPWSKTGGGLGVDVTGGLPIELVKGRHVRMTIAPRYDQYADAWDTSVVDI- 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 82 MNVVFVGAEMAPWSKTGGGLGVDVGGVPPMAAANGHRVMVSPRYDQYKDAWDTSVVSEIK 141

QY 61 MG---EKVRYPHSTKKGVRHVRWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKRFALFC 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 142 MGDGYETVRFHCHYKRGVDRVFDHPLFLERVWGKTEEKIYGPDAQYDQNLQRLSLLC 201

QY 118 KAATEAARVL-----PF--GP-GEDCVFVANDHNSALVPVLLKDEYQPKGQFTKAKSVLA 169
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 202 QAALAEAPRILSLNNPFSFGYGEDVVFVNCNDWHTGPLSCYLSKNSYQNSGIYDKDAKTAFC 261

QY 170 IHNTAFQGRMWEAEKDTKLPOAAFDKLFASDGYAKVYTEATPMEDEKPLTGKTYKKI 229
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 262 IHNISYQGRFAFSDPPELNLPERFKSSDFDIDGYEK-----PVEG---RKI 304

QY 230 NWLKGGLIADKLVTVSPNYATEITAADAAGVVELDTVIRAKGIEGIVNGMDIEEWNPKTD 289
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 305 NNMKAGILEADRVLTVPYAEELISGIARGCELDNIMRLTGITGIVNGMDVSEWDPSKD 364

QY 290 KFLSAPYDQNSVYAGKAAAEALQAEGLPVDPTAPLFAFGRLEEQKGVDIILAAALPKI 349
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 365 KYIAVKYDVSTAVEAKALKEALQAEVGLPVDNRNIPLVAFTGRLEEQKGPVMAAAIP-L 423

QY 350 LAUPKVOITAILGTGKAAYEKLVAIGTKYGRAGVVKFSAPLAHMLTAGADEMLVPSRF 409
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 424 LMEEDIQIVLLGTGKKFERMLMSAEKYPDKVRAVVKFNAALAHHIMAGADILLAVTSRF 483

QY 410 EPCGLIOLHAMHYGTVFVPAVSTGGLVDVT 438
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 484 EPCGLIOLQMRVGTGTCACASTGGLVDVTI 512
```

RESULT 11

S07314
glycogen(starch) synthase (EC 2.4.1.11) precursor - maize
N;Alternate names: starch synthase; UDP-glucose starch glycosyltransferase, starch granu
C;Species: Zea mays (maize)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 02-Aug-2002
C;Accession: S07314
R;Kloesgen, R.B.; Gierl, A.; Schwarz-Sommer, Z.; Saedler, H.
Mol. Gen. Genet. 203, 237-244, 1986
A;Title: Molecular analysis of the waxy locus of Zea mays.
A;Reference number: S07314

A;Accession: S07314
A;Molecule type: DNA
A;Residues: 1-605 <KIO>
A;Cross-references: EMBL:X03935; NID:g22509; PIDN:CAA27574.1; PID:g1644339
A;Experimental source: line C
A;Note: translation of the nucleotide sequence is not complete
A;Note: part of this sequence, including the amino end of the mature protein, was confir
C;Genetics:

A;Gene: waxy
A;Introns: 107/3; 134/3; 167/3; 197/3; 219/1; 252/3; 289/2; 370/3; 431/3; 495/3; 524/3;
C;Function:
A;Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producing e
C;Superfamily: starch synthase
C;Keywords: amyloplast; chloroplast; glycogen/starch biosynthesis; glycosyltransferase;
F;1-72/Domain: transit peptide (amyloplast) #status predicted <TMP>
F;73-605/Product: glycogen (starch) synthase #status experimental <MAT>

Query Match 51.8%; Score 1185; DB 1; Length 605;
Best Local Similarity 52.0%; Pred. No. 7.1e-77;
Matches 234; Conservative 67; Mismatches 119; Indels 30; Gaps 8;

```
QY 2 LDIVMAAEVAPWSKTGGGLGVDVTGGLPIELVKGRHVRMTIAPRYDQYADAWDTSVVDI- 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 77 MNVVFVGAEMAPWSKTGGGLGVDVGGVPPMAAANGHRVMVSPRYDQYKDAWDTSVVSEIK 136

QY 61 MG---EKVRYPHSTKKGVRHVRWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKRFALFC 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 137 MGDGYETVRFHCHYKRGVDRVFDHPLFLERVWGKTEEKIYGPVAGIDYRDNQRLSLLC 196

QY 118 KAATEAARVL-----PF--GP-GEDCVFVANDHNSALVPVLLKDEYQPKGQFTKAKSVLA 169
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 197 QAALAEAPRILSLNNPFSFGYGEDVVFVNCNDWHTGPLSCYLSKNSYQNSGIYRDAKTAFC 256

QY 170 IHNTAFQGRMWEAEKDTKLPOAAFDKLFASDGYAKVYTEATPMEDEKPLTGKTYKKI 229
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 257 IHNISYQGRFAFSDPPELNLPERFKSSDFDIDGYEK-----PVEG---RKI 299

QY 230 NWLKGGLIADKLVTVSPNYATEITAADAAGVVELDTVIRAKGIEGIVNGMDIEEWNPKTD 289
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 300 NNMKAGILEADRVLTVPYAEELISGIARGCELDNIMRLTGITGIVNGMDVSEWDPSRD 359

QY 290 KFLSAPYDQNSVYAGKAAAEALQAEGLPVDPTAPLFAFGRLEEQKGVDIILAAALPKI 349
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 360 KYIAVKYDVSTAVEAKALKEALQAEVGLPVDNRNIPLVAFTGRLEEQKGPVMAAAIPQL 419

QY 350 L-ATPKVOITAILGTGKAAYEKLVAIGTKYGRAGVVKFSAPLAHMLTAGADEMLVPSR 408
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 420 MEMVEDVQIVLLGTGKKFERMLMSAEKFPKRVAVVKFNAALAHHIMAGADVLAVTSR 479

QY 409 EPCGLIOLHAMHYGTVFVPAVSTGGLVDVT 438
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 480 EPCGLIOLQMRVGTGTCACASTGGLVDVTI 509
```

RESULT 12

T07668
starch synthase (EC 2.4.1.21) SSI precursor - potato

C;Species: Solanum tuberosum (potato)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2002
C;Accession: T07668
R;Abel, G.J.W.

submitted to the EMBL Data Library, January 1997

A;Reference number: Z16081

A;Accession: T07668

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-641 <ABE>

A;Cross-references: EMBL:Y10416

A;Experimental source: cv. Desiree; leaf

C;Genetics:

A;Gene: SSI

C;Function:

A;Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producin

A;Pathway: starch synthase

C;Superfamily: starch synthase

C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 38.0%; Score 870.5; DB 2; Length 641;
Best Local Similarity 42.9%; Pred. No. 2.3e-54;
Matches 194; Conservative 80; Mismatches 139; Indels 39; Gaps 14;

```
QY 3 DIVVMAAEVAPWSKTGGGLGVDVTGGLPIELVKGRHVRMTIAPRY-----DQYADAWDTS 55
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 132 NIIFVTAAAPYKSTGGGLGVDVGGVPPMAAANGHRVMVSPRYLNGGSDSKYANAVDLD 191

QY 56 V--VVDIMG--EKVRYFHSIKKGVRHVRWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRK 111
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 192 VRAIVCHCGDAGVAFYHEYRAGVDVFDHSSY-----RPGTF-YGDIYGA-FGDNQF 244

QY 112 RFALFCKAAIEAARVLFPGP-----GEDCVFVANDHNSALVPVLLKDEYQPKGQFTKAKSVL 168
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 245 RFTLLSHAACEAPLVLPGLGGFTYGEKCLFLANDWHAALVPLLLAAKRYGVYGVKARSTV 304
```


Job time : 16.3212 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:50:41 ; Search time 5.8117 Seconds
(without alignments)
3088.950 Million cell updates/sec

Title: US-09-980-771a-7

Perfect score: 2288

Sequence: 1 ALDIVMAAEVAPWSKTGGL.....AMHYGTVPVVASTGGLVDVTV 438

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1301	56.9	608	1	UGST_MANES
2	1277	55.8	608	1	UGST_ANTMA
3	1275	55.7	607	1	UGST_SOLTU
4	1258	55.0	603	1	UGST_PEA
5	1246	54.5	608	1	UGST_IPOBA
6	1232.5	53.9	603	1	UGST_HORVU
7	1229	53.7	609	1	UGST_ORVGL
8	1229	53.7	609	1	UGST_ORVSA
9	1204	52.6	615	1	UGST_WHEAT
10	1187	51.9	608	1	UGST_SORBI
11	1185	51.8	605	1	UGST_MAIZE
12	870.5	38.0	641	1	UGS2_SOLTU
13	870	38.0	626	1	UGS2_ORVSA
14	850	37.2	610	1	UGS2_WHEAT
15	833	36.4	752	1	UGS3_PEA
16	805.5	35.2	788	1	UGS3_SOLTU
17	660.5	28.9	477	1	GLGA_STRPN
18	638.5	27.5	485	1	GLGA_BACST
19	628	27.4	484	1	GLGA_BACST
20	623.5	27.3	480	1	GLGA_AGR5
21	615	26.9	480	1	GLGL_RHIME
22	611	26.7	486	1	GLGA_THEMA
23	605.5	26.5	480	1	GLGA_RHTR
24	601	26.3	477	1	GLGA_CLOMB
25	601	26.3	478	1	GLGA_LACIA
26	589	25.7	484	1	GLGA_VIBCH
27	587.5	25.7	482	1	GLGA_CLOPE
28	581	25.4	476	1	GLGA_VERPE
29	578.5	25.3	476	1	GLGA_BACHD
30	575.5	25.2	481	1	GLGA_RHILLO
31	570.5	24.9	486	1	GLG2_RHIME
32	566	24.7	477	1	GLGA_SALTI
33	565	24.7	477	1	GLGA_ECOLI

34	559	24.4	477	1	GLGA_SALTY	P05416 salmonella
35	552	24.1	480	1	GLGA_PASMU	Q9cn91 pasteurella
36	528	23.1	1230	1	UGS4_SOLTU	Q43846 solanum tub
37	527.5	23.1	465	1	GLGA_SYNP7	Q93577 synechococc
38	524	22.9	476	1	GLGA_HAEIN	P45179 haemophilus
39	522.5	22.8	463	1	GLGA_AOUAE	O66935 aquifex aeo
40	520.5	22.7	472	1	GLGA_ANASP	Q8yvu5 anabaena sp
41	514	22.5	477	1	GLGA_RHOSH	Q9rrh6 rhodobacter
42	506.5	22.1	477	1	GLGA_SYNY3	P74521 synechocyst
43	502	21.9	492	1	GLG2_ANASP	Q8z0g9 anabaena sp
44	496.5	21.7	491	1	GLG2_SYNY3	P76233 synechocyst
45	493.5	21.6	461	1	GLGA_FUSNN	Q8rrf65 fusobacteri

ALIGNMENTS

RESULT 1

ID	UGST_MANES	STANDARD;	PRT;	608 AA.
AC	Q43784;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Granule-bound glycogen [starch] synthase, chloroplast precursor (EC 2.4.1.11).			
GN	WAXY OR GBSS.			
OS	Manihot esculenta (Cassava) (Manioc).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid 1; Malpighiales; Euphorbiaceae; Manihot.			
OX	NCBI_TaxID=3983;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. M.COI.22; TISSUE=Tuberous root;			
RX	MEDLINE=94083565; PubMed=8260633;			
RA	Salchuzman S.N., Jacobsen E., Visser R.G.F.;			
RT	"Isolation and characterization of a cDNA encoding granule-bound starch synthase in cassava (Manihot esculenta Crantz) and its antisense expression in potato.";			
RT	Plant Mol. Biol. 23:947-962(1993).			
RL	CC -!- FUNCTION: RESPONSIBLE FOR THE SYNTHESIS OF AMYLOSE IN RESERVE STARCH.			
CC	CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-alpha-D-glucosyl}(N) = UDP + {(1,4)-alpha-D-glucosyl}(N+1).			
CC	CC -!- PATHWAY: Starch biosynthesis.			
CC	CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.			
CC	CC -!- TISSUE SPECIFICITY: SYNTHESIZED IN A NUMBER OF DIFFERENT ORGANS, BUT MOST ABUNDANTLY IN TUBERS.			
CC	CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE FAMILY.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC	EMBL; X74160; CAA52273.1; .
DR	InterPro; IPR001296; Glycos_transf_1.
DR	Pfam; PF00534; Glycos_transf_1; 1.
KW	Glycogen biosynthesis; Transferease; glycosyltransferase;
KW	Transit peptide; Chloroplast; Starch biosynthesis.
FT	TRANSIT 1 78 CHLOROPLAST (BY SIMILARITY).
FT	CHAIN 79 608 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.
FT	BINDING 96 96 UDP-GLUCOSE (BY SIMILARITY).
SQ	SEQUENCE 608 AA; 66968 MW; C9C970CD3011BDD8 CRC64;

Query Match 56.9%; Score 1301; DB 1; Length 608;
Best Local Similarity 57.5%; Pred. No. 1.3e-85;
Matches 258; Conservative 51; Mismatches 110; Indels 30; Gaps 7;

*Sequence of the structural gene for granule-bound starch synthase of
RT potato (Solanum tuberosum L.) and evidence for a single point
RT deletion in the amf allele. ;
RL Mol. Gen. Genet. 228:240-248(1991).
RN [2]

RP REVISIONS.

RA van der Lei J.F.R.;

RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-cv. Dongnong 303;

RA Dai W.L., Deng W., Cui M., Xiu M., Zhao S.Y., Wang X.M.;

RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.

CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-alpha-D-glucosyl}(N) =

CC UDP + {(1,4)-alpha-D-glucosyl}(N+1).

CC -!- PATHWAY: Starch biosynthesis.

CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.

CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE

CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X58453; CAA41359.1; -;

DR EMBL; X83220; CAA58220.1; -;

DR PIR; S16555; YUPOY.

DR InterPro; IPR001296; Glycos_transf_1.

DR Pfam; PF00534; Glycos_transf_1; 1.

KW Glycogen biosynthesis; Transferase; Glycosyltransferase;

KW Transit peptide; Chloroplast; Starch biosynthesis.

FT TRANSIT 1 77 CHLOROPLAST.

FT CHAIN 78 607 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.

FT BINDING 95 95 UDP-GLUCOSE (BY SIMILARITY).

FT CONFLICT 130 130 A -> T (IN REF. 3).

FT CONFLICT 398 398 I -> V (IN REF. 3).

FT SEQUENCE 607 AA; 66575 MW; 2A377865CEFA650 CRC64;

Query Match 55.0%; Score 1275; DB 1; Length 607;

Best Local Similarity 56.3%; Pred. No. 9.6e-84;

Matches 253; Conservative 54; Mismatches 112; Indels 30; Gaps 7;

QY 2 LDIVMVAEAPWSKTGGGLDVTGGPLPIELVKRHRVMTIAPRYDQYADAWDTSVVVDI- 60

DB 81 MNLIFVGTGVEGVPWSKTGGGLDVTGGPLPALAARGHRVMTISPRYDQYKDAWDTSVAVEVK 140

QY 61 MG---EKVYRTHSIKKGVRHWIDHPWFLAKVWGKTSKLYGPRSGADYLDNKKRFAIFC 117

DB 141 VGDSTEIVRFFHCYKRGVDRVFDHPHMELEKVGWGTGSKIYGPAGLDYLDNLRFSLLC 200

QY 118 KAAIEAARVLFP-----GP-GEPCVFVANDHWSALVPVLLKDEYQPKGQFTAKKSVAL 169

DB 201 QAALEAPKVLNLSNYSYSGPGEVDYFIANDWHHTALIPCYLKSMSYQSGIYLNKAVFC 260

QY 170 IHNIAFGQRMWEAPKTKLPQAAFKLAISDGYAKVYTEATPMEDEKPLTGTGYKKI 229

DB 261 IHNIAFGQRFSDPFLNLEDFRGSFDFIDGYEK-----PVKG---RKI 303

QY 230 NWLKGSIITADKLVTVSPNYATEIATAADAGGVVELDTVIRAKGIEIVNGMDIEENPKTD 289

DB 304 NWMKAGILSHRVTVSPYIAQELVSAYDKGVGLDSVLKTKITGIVNGMDTQEWNPATD 363

QY 290 KFLSAPYDQNSYAGKAAKALQALQELGLPVDPTAPLFAFTGRLEEQGVLDIILALPKI 349

DB 364 KYTDVKYDITTVMDAKPLKALQAAVGLPDKTKIPLIGFGRLEEQGSDITLVAIHKF 423

QY 350 LATPKVQITAILGTGKAAVEKLVNALIGTKYKGRAGVWKFSAPLAHMLTAGADFMVPSRF 409

DB 424 IGL-DVQIVLVTGKKEFEQETQLEVLVYPNKAQVAFNVLPHMITAGADFMVPSRF 482

QY 410 EPGQLQLHAMHYGTVPVWASTGGLVDTV 438

DB 483 EPGQLQLHAMRYGTVPICASTGGLVDTV 511

RESULT 4

UGST_PEA

ID UGST_PEA

AC 043092;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Granule-bound glycogen [starch] synthase, chloroplast precursor

DE (EC 2.4.1.11) (CBSSI).

OS Pisum sativum (Garden pea).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.

OC NCBI_TaxID=3886;

OX [1]

RN SEQUENCE FROM N.A., AND SEQUENCE OF 76-88.

RC STRAIN-cv. BC1/RR; TISSUE=Embryo;

RX MEDLINE=93251108; PubMed=1302049;

RA Dry I., Smith A., Edwards A., Bhattacharyya B., Dunn P., Martin C.;

RT "Characterization of cDNAs encoding two isoforms of granule-bound

RT starch synthase which show differential expression in developing

RT storage organs of pea and potato.;"

RL Plant J. 2:193-202(1992).

CC -!- FUNCTION: MAY BE RESPONSIBLE FOR THE SYNTHESIS OF AMYLOSE.

CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-alpha-D-glucosyl}(N) =

CC UDP + {(1,4)-alpha-D-glucosyl}(N+1).

CC -!- PATHWAY: Starch biosynthesis.

CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.

CC -!- TISSUE SPECIFICITY: EXPRESSED IN PODS AND LEAVES. NO EXPRESSION

CC IN FLOWERS OR STIPULES.

CC -!- DEVELOPMENTAL STAGE: EXPRESSED AT ALL STAGES OF EMBRYONIC

CC DEVELOPMENT WITH HIGHEST LEVELS IN LATER DEVELOPMENTAL STAGES.

CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE

CC FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X88789; CAA61268.1; -;

DR InterPro; IPR001296; Glycos_transf_1.

DR Pfam; PF00534; Glycos_transf_1; 1.

KW Glycogen biosynthesis; Transferase; Glycosyltransferase;

KW Transit peptide; Chloroplast; Starch biosynthesis.

FT TRANSIT 1 75 CHLOROPLAST.

FT CHAIN 76 603 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.

FT BINDING 91 91 UDP-GLUCOSE (BY SIMILARITY).

FT SEQUENCE 603 AA; 66362 MW; 817252FDD12CCAA0 CRC64;

Query Match 55.0%; Score 1258; DB 1; Length 603;

Best Local Similarity 54.9%; Pred. No. 1.6e-82;

Matches 248; Conservative 63; Mismatches 105; Indels 36; Gaps 9;

QY 2 LDIVMVAEAPWSKTGGGLDVTGGPLPIELVKRHRVMTIAPRYDQYADAWDTSVVVDI- 60

DB 77 MSLVFGVEGVPWSKTGGGLDVTGGPLVLAAGNHRVMTVSPRYDQYKDAWDTNVLVEVK 136

QY 61 MGEK---VRYFHSIKKGVRHWIDHPWFLAKVWGKTSKLYGPRSGADYLDNKKRFAIFC 117

DB 137 VGDKTETVRFHCYKRGVDRVFDHPHMELEKVGWGTGSKIYGPAGLDYLDNLRFSLLC 196

QY 118 KAAIEAARVLFP-----GP-GEPCVFVANDHWSALVPVLLKDEYQPKGQFTAKKSVAL 169

Db 197 QAALEAPRVNLNSSKFGSGYGEDVLFVANDWHSALIPCVLKNYSRGLYKNAKVAFC 256
 QY 170 IHNTAFQGRWEEAFK--TKLPQAFKLAESDGYAKVYTEATPMEDEKPLTKTKY 226
 Db 257 IHNTAYOGR--NAFSDLSLUNPDEPSSDFIDGYNK-----PCEG--- 296
 QY 227 KKNLWKGIIAADKLVTPSPNATEIAADAAGVELDTVIRAKGIEGIVNGMDIEEWN 286
 Db 297 KKNWMAKAGILESDQVETVSPHYAKELISGEDRGVLDNIRSTGIIGVNGMDREWS 356
 QY 287 KTKFELSPADQNSVYAGKAAAEALQAEGLPVDPTAPFAFGRLEEQKGVDIILAA 346
 Db 357 QTDRIEYDHYNETTVEAKPLKGLTQAEGLPVDSSPLIGFGRLEEQKGSILVEAI 416
 QY 347 PKILATPKVOIATIGTCKAAVEKLVNAIGTKYGRAGVYKFSAPLAHMLFAGADEMLVP 406
 Db 417 AK-FADENVQIVLGTGCKKMEKQIEVEKYPGKAIGTKFNSPLAKHIIAGADFVIP 475
 QY 407 SRFEPCGLIOLHAMHYGTVPVASTGGGLVDIV 438
 Db 476 SRFEPCGLVOLHAMPYGTVPVLSVSTGGGLVDIV 507

RESULT 5

UGST_IPOBA STANDARD; PRT; 608 AA.
 AC Q42857;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Granule-bound glycogen [starch] synthase, chloroplast precursor
 DE (EC 2.4.1.11).
 GN WAXY OR SS67.
 OS Ipomoea batatas (Sweet potato) (Batace).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Convolvulaceae; Ipomoea.
 OX NCBI_TaxID=4120;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Tainong 57; TISSUE=Tuberous root;
 RA Wang S.J., Yeh K.W., Tsai C.Y.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYLOSE IN ENDOSPERM.
 CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-alpha-D-glucosyl}(N) =
 CC UDP + {(1,4)-alpha-D-glucosyl}(N+1).
 CC -!- PATHWAY: Starch biosynthesis.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U4126; AAA86423.1; -;
 CC InterPro: IPR001296; Glycos_transf_1.
 CC Pfam: PF00534; Glycos_transf_1;
 CC KW Glycogen biosynthesis; Transferease; Glycosyltransferase;
 CC KW Transit peptide; Chloroplast; Starch biosynthesis.
 CC TRANSIT 1 76 CHLOROPLAST (BY SIMILARITY).
 CC CHAIN 77 608 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.
 CC BINDING 96 96 UDP-GLUCOSE (BY SIMILARITY).
 CC SEQUENCE 608 AA; 67000 MW; CCA8FA50A2F69CB0 CRC64;
 CC -----

Query Match 54.5%; Score 1246; DB 1; Length 608;
 Best Local Similarity 56.3%; Pred. No. 1.1e-81;
 Matches 253; Conservative 48; Mismatches 118; Indels 30; Gaps 7;

QY 2 LDIVMAAEVAPWSKTGGGLDVTGGLPIELVKRGRHVRMTIAPRYDQYADAMDTSVVVD-I 60
 Db 82 MNLVFGCEEGPCKTGGGLDVLGGLPPALAAARGHRVMTVCPRYDQYKDAWETCVVVEPQ 141
 QY 61 MG---EKVRYFHSIKKGVHRVWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHKBFALFC 117
 Db 142 VGDRIEYRFRFHSYKRGVDRVFDHPFLEKVMGKTSMLYGPAGKDYKDNQLRFLSLC 201
 QY 118 KAAETAAARVLPF-----GP-GEPCVFVANDWHSALVPVLLKDEYQPKGQFTKAKSVLA 169
 Db 202 QAALEAPRVNLNSSKFGSGYGEDVVFVANDWHSALPCLTKMYQSRGIYMNKAVAF 261
 QY 170 IHNTAFQGRWEEAFKTKLPQAFKLAESDGYAKVYTEATPMEDEKPLTKTKYKKI 229
 Db 262 IHNTAYOGRFAFSDFSLLNLPDEYKGSDFIDGYNK-----PVKG---RKI 304
 QY 230 NNLKGGIIAADKLVTPSPNATEIAADAAGVELDTVIRAKGIEGIVNGMDIEEWNPKTD 289
 Db 305 NNMKAGITREADRVTVPNYAKELVSCVKGVELDNHINDGICITGCMGMDQENWPAID 364
 QY 290 KFLSAPYDQNSVYAGKAAAEALQAEGLPVDPTAPFAFGRLEEQKGVDIILAAALPKI 349
 Db 365 KYLAVKYDITVVMQAKPLKALQAAVGLPVDNRNPLIGFGRLEEQKGSIDLIAAISKF 424
 QY 350 LATPKVQIATIGTCKAAVEKLVNAIGTKYGRAGVYKFSAPLAHMLFAGADFMLVPSRF 409
 Db 425 ISM-DVQITLILGTGKKFEQQLQEVNMYPKARGVAKENVPLAHMITAGADFMLIPSRF 483
 QY 410 EPCGLIOLHAMHYGTVPVASTGGGLVDIV 438
 Db 484 EPCGLIOLHAMRYGTPICASTGGGLVDIV 512

RESULT 6

UGST_HORVU STANDARD; PRT; 603 AA.
 AC P09842;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Granule-bound glycogen [starch] synthase, chloroplast precursor
 DE (EC 2.4.1.11).
 GN WAXY.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;
 OC Triticeae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Vogelsanger Gold; TISSUE=Leaf;
 RX MEDLINE=88303345; PubMed=2970062;
 RA Rhode W., Becker D., Salamini F.;
 RT "Structural analysis of the waxy locus from Hordeum vulgare.";
 RL Nucleic Acids Res. 16:7185-7186(1988).
 RN [2]
 RP SEQUENCE OF 76-89.
 RC STRAIN=cv. H354-295-2-5; TISSUE=Starchy endosperm;
 RX MEDLINE=94170739; PubMed=8125056;
 RA Flengsrud R.;
 RT "Separation of acidic barley endosperm proteins by two-dimensional
 RT electrophoresis.";
 RL Electrophoresis 14:1060-1066(1993).
 CC -!- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYLOSE IN ENDOSPERM.
 CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-alpha-D-glucosyl}(N) =
 CC UDP + {(1,4)-alpha-D-glucosyl}(N+1).
 CC -!- PATHWAY: Starch biosynthesis.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.
 CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the Swiss Institute of Bioinformatics and the EMBL outstation -


```

Db 484 EPCGLIQGMRYTGPCACASTGGLVDTI 512

RESULT 11
UGST_MAIZE STANDARD; PRT; 605 AA.
ID P04713;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Granule-bound glycogen [starch] synthase, chloroplast precursor
DE (EC 2.4.1.11).
GN WAXY.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
AC Kioesgen R.B., Gierl A., Schwarz-Sommer Z., Saedler H.;
RT "Molecular analysis of the waxy locus of Zea mays.";
RL Mol. Gen. Genet. 203:237-244(1986).
CC -!- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYLOSE IN ENDOSPERM.
CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-alpha-D-glucosyl}(N) =
CC UDP + {(1,4)-alpha-D-glucosyl}(N+1).
CC -!- PATHWAY: starch biosynthesis.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X03935; CAA27574.1; -.
DR EMBL; M24258; AAA33520.1; -.
DR PIR; S07314; S07314.
DR MaizeDB; 15806; -.
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycos_transf_1;
KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW Transit peptide; Chloroplast; Starch biosynthesis.
FT TRANSIT 1 72 CHLOROPLAST.
FT CHAIN 73 605 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.
FT BINDING 91 91 UDP-GLUCOSE (BY SIMILARITY).
SQ SEQUENCE 605 AA; 65966 MW; 137F15207DBFC189 CRC64;

Query Match 51.8%; Score 1185; DB 1; Length 605;
Best Local Similarity 52.0%; Pred. No. 2.6e-77;
Matches 234; Conservative 67; Mismatches 119; Indels 30; Gaps 8;

QY 2 LDIWVAEAPWSKTSGLGVDVTCGLPIELVKRGHRVMTIAPRYDQYADAWDTSVYVDI- 60
Db 77 MNVVFVGAEMAPWSKTSGLGVDVTCGLPMAANGHRVWVSPRYDQYKDAWDTSVVSEIK 136
QY 61 MG---EKVRYFHSIKKGVHVRWIDHPWFLAKVWKTGSKLYGPRSGADYLNHKKFALFC 117
Db 137 MGDYEIVRFPHCYKRGVDRVFDVDPFLFLEKRVGKTEKIKYGPVAGTDYRDNQRFSLLC 196
QY 118 KAAEAAARVL-----PF--GDCVFVANDWHSALVPVLLKDEYQPKGQFTAKSVLA 169
Db 197 QAALAEAPRLSLNNPFSGPGYEDVFCVNDWHTGPLSCYLKSNYQSHGIYRDAKATFC 256
QY 170 IHNIAFGQRMWEAFKTKLPQAADFCLAFSDGAKVYTEATPMEEDKPLTKYTKYKI 229
Db 257 IHNISYQGRFAFSYDPNLNLPFRPKSSFDYIDGYEK-----PVEG---RKI 299
QY 230 NWLKGGLIAADKLVTSPNYPATELAADAAGGVGLDVTIRAKGIEGIVNGMDIEENPKTD 289

```

```

Db 300 NNMKAGILEADRVLTSPYIEELISGTARGCELDNIMRLTGITVINGMDYSEWDPSRD 359
QY 290 KFLSAPYDONSVYACGAARAKALQAEGLGLPVLPTAFGLRLEQKGVDIILAALPKI 349
Db 360 KYIAVKYDVSTAVEAKALNKEALQAEVGLPVRNIPLVAFGLRLEQKGPVMAAIIQL 419
QY 350 L-ATPKVQIATILGTGKAAYEKIVNAIGTKYKRAKGVVVKFSAPLAHMLTAGADEMLVPSR 408
Db 420 MEMVEDVQIVLLGTGKKKFERMLMSAEKFPCKVRVVKFNAALAHMAGADVLAITSR 479
QY 409 FEPCCGLIQHAMHYGTVPVVASTGGLVDTV 438
Db 480 FEPCCGLIQGMRYTGPCACASTGGLVDTI 509

RESULT 12
UGS2_SOLTU STANDARD; PRT; 641 AA.
ID P93568;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Soluble glycogen [starch] synthase, chloroplast precursor
DE (EC 2.4.1.11) (SS I).
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
AC STRAIN-cv. Desiree; TISSUE=Leaf;
RT MEDLINE=97164391; PubMed=9011082;
RA Abel G.J.W., Springer F., Willmitzer L., Kossmann J.;
RT "Cloning and functional analysis of a cDNA encoding a novel 139 kDa
RT starch synthase from potato (Solanum tuberosum L.).";
RL Plant J. 10:981-991(1996).
CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-alpha-D-glucosyl}(N) =
CC UDP + {(1,4)-alpha-D-glucosyl}(N+1).
CC -!- PATHWAY: starch biosynthesis.
CC -!- SUBCELLULAR LOCATION: AMYLOPLAST OR CHLOROPLAST, SOLUBLE (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y10416; CAA71442.1; -.
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycos_transf_1;
KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW Transit peptide; Chloroplast; Amyloplast; Starch biosynthesis.
FT TRANSIT 1 641 ? CHLOROPLAST (POTENTIAL).
FT CHAIN 145 145 SOLUBLE GLYCOGEN [STARCH] SYNTHASE.
FT BINDING 145 145 UDP-GLUCOSE (BY SIMILARITY).
SQ SEQUENCE 641 AA; 70608 MW; 30FB0546CEFB74C CRC64;

Query Match 38.0%; Score 870.5; DB 1; Length 641;
Best Local Similarity 42.9%; Pred. No. 8.2e-55;
Matches 194; Conservative 80; Mismatches 139; Indels 39; Gaps 14;

QY 3 DIWVAAEAPWSKTSGLGVDVTCGLPIELVKRGHRVMTIAPRY-----DOYADAWDT 55
Db 132 NIIFVTAAAPYSKTSGGLGVDVTCGLPMAALAAKGRVWVSPRYNGGSPDEKYANAVDLD 191
QY 56 V--VVDIMG--EKVRYFHSIKKGVHVRWIDHPWFLAKVWKTGSKLYGPRSGADYLDNHK 111

```

Db 192 VRATVHCFGDAGEVAFYHEIRAGVDWVFDHSSYC-----RPGTP-YGDIYA-FGDNOF 244
 QY 112 RFALPKCAIARVLPFGP---GEDCVFVANDHWSALVPVLLKDEYOPKGFTRAKSVL 168
 Db 245 RFTLLSHAAACEAPLPLGGFTYGEKCLFLANDHAALVPLLLAAKYPGYVKDARSIV 304
 QY 169 AIHNIAGFGRWEEAFKOTKLPQAAFDKLAFSDGAKVYVTEATPDEEKEKPLTKTKYK 228
 Db 305 AIHNIAGFVEPVTYNNIGLPPQWYGAV---EWIFPTWARAHALD-----TGET--- 351
 QY 229 INWLKGGIIAADKLVTSPNYATEIAADAAGVDELDTVI--RAKIGIEIVNGMDIEWNPK 286
 Db 352 VNLKGAIAVADRILTVSGYSWEITP-PEGYGLHELLSSQSVLNGITNGIDVNDWP 410
 QY 287 KTDKELSPYQNSVYAGKAAKAEALQAEGLPVDPTAPLFAFICRLEEQKGVDIILAL 346
 Db 411 STDEHIAHSYINDL-SGVQCKTDLOKELGLPIRDCPLIGFICRLDYQKGVDIILSAI 469
 QY 347 PKILATPKVQIAILGTGKAAEKLVAIGTKYGRKGVKFSAPLAHLMTAGADFMLVP 406
 Db 470 PELMQN-DVQVVLGSGEKQVEDMWRHTENLFDKFRWGVNFPVSHRITAGCDLILMP 528
 QY 407 SRFEPCGLIQLHAMHYGTVPVVAATGGLVDIV 438
 Db 529 SRFEPCGLNQLYAMRYGTPVHSIGGLRDIV 560
 RESULT 13
 UGS2_ORYZA
 ID UGS2_ORYZA STANDARD; PRT; 626 AA.
 AC Q40739;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Soluble glycogen [starch] synthase, chloroplast precursor
 DE (EC 2.4.1.11) (SSS)
 OS Oryza sativa (Rice)
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 114-131.
 RC STRAIN=cv, Japonica; TISSUE=Seed;
 RX MEDLINE=94302151; PubMed=7518089;
 RA Baba T., Nishihara M., Mizuno K., Kawasaki T., Shimada H.,
 RA Kobayashi E., Omishi S., Tanaka K.-I., Arai Y.;
 RT "Identification, cDNA cloning, and gene expression of soluble starch
 RT synthase in rice (Oryza sativa L.) immature seeds.";
 RL Plant Physiol. 103:565-573(1993).
 CC -!- CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-alpha-D-glucosyl)(N) =
 CC UDP + ((1,4)-alpha-D-glucosyl)(N+1).
 CC -!- PATHWAY: starch biosynthesis.
 CC -!- SUBCELLULAR LOCATION: AMYLOPLAST OR CHLOROPLAST, SOLUBLE.
 CC -!- TISSUE SPECIFICITY: LEAVES AND IMMATURE SEEDS.
 CC -!- MISCELLANEOUS: THREE FORMS OF SOLUBLE STARCH SYNTHASE WERE
 CC PURIFIED: RSS1, RSS2 AND RSS3.
 CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
 CC FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D16202; BAA03739.1; -
 CC InterPro; IPR001296; Glycos_transf_1.
 CC Pfam; PF00534; Glycos.transf.1; 1.
 KW Glycogen biosynthesis; Transferase; Glycosyltransferase;

KW Transit peptide; Chloroplast; Amyloplast; Starch biosynthesis.
 FT TRANSIT 1 113 CHLOROPLAST (POTENTIAL).
 FT CHAIN 114 626 SOLUBLE GLYCOGEN [STARCH] SYNTHASE, RSS3.
 FT CHAIN 122 626 SOLUBLE GLYCOGEN [STARCH] SYNTHASE, RSS1.
 FT BINDING 147 147 UDP-GLUCOSE (BY SIMILARITY).
 SQ SEQUENCE 626 AA; 68451 MW; 03E4182507D26658 CRC64;
 Query Match 38.0%; Score 870; DB 1; Length 626;
 Best Local Similarity 42.7%; Pred. No. 8.6e-55;
 Matches 192; Conservative 74; Mismatches 146; Indels 38; Gaps 12;
 QY 4 IVNVAEAVAPSKTGGIGDVTGGLPIELVKRHRVMTIAPRY-----DOYADANDTSVV 57
 Db 135 VVFTVGEAPSKYAGSGIGDVGSLPIALALRHRVMMVMPRYMNGALNKNFANAFYTERK 194
 QY 58 VDIM---GE-KVRYFHSIKKGVHRVWIDHPWELAKVWGKTSKLYGPRSGADYLDNHKRF 113
 Db 195 IKIPCFGEHEVTFEYRDSVDWVVDHPSY-----HRPNLYGDNFGA-FGDNQFRY 247
 QY 114 ALFCKAAIEAARVLPFGP---GEDCVFVANDHWSALVPVLLKDEYOPKGFTRAKSVLAI 170
 Db 248 TLLCYAAACEAPLILELGGVYIGKCMFVVDHWSALVPVLLAAKYRPGYVYRDARSVLI 307
 QY 171 HNIAGFGRWEEAFKOTKLPQAAFDKLAFSDGAKVYVTEATPDEEKEKPLTKTKYKKN 230
 Db 308 HNLAHOGVEPASTYDPLGLPPEWYGALEW-----VFPEWARRHADKG-----EAVN 354
 QY 231 WLKGGIIAADKLVTSPNYATEIAADAAGVDELDTVI--RAKIGIEIVNGMDIEWNPKT 288
 Db 355 FLKGAVVTADRIVTVSQYSWEVTT-AGGGQGLNELLSRKSVLNGIVNGIDINDWNPST 413
 QY 289 DKFLSAPYQNSVYAGKAAKAEALQAEGLPVDPTAPLFAFICRLEEQKGVDIILALAPK 348
 Db 414 DKFLPYHSVDDL-SGAKCAELQKELGLPIRDPVPLIGFICRLDYQKGLDIKLAIPD 472
 QY 349 ILATPKVQIAILGTGKAAEKLVAIGTKYGRKGVKFSAPLAHLMTAGADFMLVPSR 408
 Db 473 LM-RDNIQFVLMGSDGDFEGMWRSTESGDKFRGVGVFVSHRITAGCDLILMPSR 531
 QY 409 FEPCLQLIQLHAMHYGTVPVVAATGGLVDIV 438
 Db 532 FEPCLNQLYAMQYGTVPVHGTGGLRDIV 561
 RESULT 14
 UGS2_WHEAT
 ID UGS2_WHEAT STANDARD; PRT; 610 AA.
 AC Q43654;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Soluble glycogen [starch] synthase, chloroplast precursor
 DE (EC 2.4.1.11) (Fragment).
 OS Triticum aestivum (Wheat).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 CC Triticeae; Triticum.
 OX NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv, T.A. Florida; TISSUE=Endosperm;
 RA Block M., Loez H., Luetticke S.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-alpha-D-glucosyl)(N) =
 CC UDP + ((1,4)-alpha-D-glucosyl)(N+1).
 CC -!- PATHWAY: starch biosynthesis.
 CC -!- SUBCELLULAR LOCATION: AMYLOPLAST OR CHLOROPLAST, SOLUBLE (BY
 CC SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
 CC FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U48227; RAB02197.1; -;
 DR InterPro: IPR001296; Glycos_transf_1.
 DR Pfam: PF00534; Glycos_transf_1.
 KW Glycogen biosynthesis; Transferrase; Glycosyltransferase;
 KW Transit peptide; Chloroplast; Amyloplast; Starch biosynthesis.
 FT NON_TER 1
 FT TRANSIT 1
 FT CHAIN 1
 FT BINDING 7
 FT SEQUENCE 610 AA; 67143 MW; 46080A3B7F87193 CRC64;

Query Match 37.2%; Score 850; DB 1; Length 610;
 Best Local Similarity 43.0%; Pred. No. 2.2e-53;
 Matches 191; Conservative 71; Mismatches 144; Indels 38; Gaps 12;

QY 10 EVAPWSKTGGLGVDVTGLPTLVKGRHVMVMTIAPRY-----DOYADAWDTSVVVDIM-- 61
 DB 1 EAAPYAKSGGLGVDVCGSLPTALAAARGHRVMVMPRYLNGSSDKNYAKALYAKHIKIPCF 60
 QY 62 --GEKVRYPHSIKKGVRHWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKFAFCKA 119
 DB 61 GGSHEVTFPFEYRDNVDVFDHP-----SYHRPGS-LYGDNFGA-FGDNQFRYLLCYA 113
 QY 120 AIEARVLPFGP---GEDCVFVANDHSAIVPVLLKDEYOPKQFTAKSVLAHNAFQ 176
 DB 114 ACEAPLLELGGYTYGNCMFVNDHSAIVPVLLAARYPYGYRDSRSTLVVHNAHQ 173
 QY 177 GRMWEAFKTKLPQAAFDKLAISDGAKVYVTEATPMEDEKPLTKTKYKINWLGKI 236
 DB 174 GVEPASTYDGLPPEWYGALEW-----VFPWARHRLDKG-----EAVNFKGAV 220
 QY 237 IADKLVTPSPNYATEIAADAAGGVLDTVI--RAKGIEGIVNGMDIEENPKTKPLISA 294
 DB 221 VTADRIYVSGQSWYETI-AGSQGLNELLSRRKSVLNGVINGIDINDWNTPTDKCLPH 279
 QY 295 PYDONSIVYAGAAKAEALQELGLPVDPTAPLFAFGRLEBEQKGVLDIILALPKLATPK 354
 DB 280 HYSYDDI-SGKAKCABELOKELGLPVREDVPLIGFGLKDYOKGIDITKAIPELM-RED 337
 QY 355 VQIAILGTGAAYEKLYNAIGTKYKRAKGVKVFSAFLAHMLTAGADFMVPSRFEPCGL 414
 DB 338 VQFVNLGSGDPFEGWRSTESSYKDKFERGWGVSVPVSHRITAGCDILLMPSRFEPCGL 397
 QY 415 IQLHAMHYGTPVVPVASTGGLVDTV 438
 DB 398 NQLYAMQYGTVPVVGTTGGLRDTV 421

RESULT 15
 UGS3_PEA
 ID UGS3_PEA STANDARD; PRT; 752 AA.
 AC Q43093;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Glycogen [starch] synthase, chloroplast precursor (EC 2.4.1.11)
 DE (GBSSI) (granule-bound starch synthase II).
 OS Pisum sativum (Garden pea).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucotsid I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
 OX NCBI_TaxID=3888;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 58-73.
 RC STRAIN=cv. BCL/RR; TISSUE=Embryo;
 RX MEDLINE=93251108; PubMed=1302049;
 RA Dry I., Smith A., Edwards A., Bhattacharyya B., Dunn P., Martin C.;

RT "Characterization of cDNAs encoding two isoforms of granule-bound
 RT starch synthase which show differential expression in developing
 RT storage organs of pea and potato.";
 RL Plant J. 2:193-202(1992).
 CC -|- CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-alpha-D-glucosyl)(N) =
 CC UDP + ((1,4)-alpha-D-glucosyl)(N+1).
 CC -|- PATHWAY: Starch biosynthesis.
 CC -|- SUBCELLULAR LOCATION: CHLOROPLAST/AMYLOPLAST, SOLUBLE AND GRANULE-
 CC BOUND.
 CC -|- TISSUE SPECIFICITY: WIDELY EXPRESSED.
 CC -|- DEVELOPMENTAL STAGE: MOST HIGHLY EXPRESSED IN EARLY EMBRYOS.
 CC LEVELS DECLINE IN LATER STAGES OF DEVELOPMENT.
 CC -|- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
 CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X89790; CAA61269.1; -;
 DR InterPro: IPR001296; Glycos_transf_1.
 DR Pfam: PF00534; Glycos_transf_1.
 KW Glycogen biosynthesis; Transferrase; Glycosyltransferase;
 KW Transit peptide; Chloroplast; Amyloplast; Starch biosynthesis.
 FT TRANSIT 1
 FT CHAIN 58
 FT BINDING 275
 FT SEQUENCE 752 AA; 83617 MW; E0496420CD359395 CRC64;

Query Match 36.4%; Score 833; DB 1; Length 752;
 Best Local Similarity 41.8%; Pred. No. 4.8e-52;
 Matches 189; Conservative 73; Mismatches 142; Indels 48; Gaps 12;

QY 2 LDIWVAEAVAPWSKTGGLGVDVTGLPTLVKGRHVMVMTIAPRYDOYADAWDTSV--VVD 59
 DB 261 MNIIIVSAECAPWSKTGGLGVDVAGSLPKALARRHVRMVAVPHYGNYAEADHIGVRYKRYK 320
 QY 60 IMGE--KVRYPHSIKKGVRHWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKFAFCK 117
 DB 321 VAGDMEVTFPFTYIDGVDTVFIIDSP-----IFRMESNYI-----GGNELDLIRWVLF 371
 QY 118 KAAIEAARVLP-----FGGEDCVFVANDHSAIVPVLLKDEYOPKQFTAKSVLAHNA 172
 DB 372 KAAVEVPWHVPCGGICYGDG-NLVPIANDHWTALLPVYLKAYVRHGLMNYTRSLVLIHN 430
 QY 173 IAFQGRWEEAFKTKLPQAAFDKLAISDGAKVYVTEATPMEDEKPLTKTKYKINWLGKI 232
 DB 431 IAHQGRG-----EVEDFTVLDLGNLDLDFKMYDPVGGEH-----FNIF 469
 QY 233 KGGIITADKLVTPSPNYATEIAADAAGGVLDTVIRAKG--IEGIVNGMDIEENPKTKD 290
 DB 470 AAGLKATADRIYVSHGYAWELKT-SEGWGLHNIINESDWKFRGIVNGVDIKWNPQFDA 528
 QY 291 FLSAP-----YDONSIVYAGAAKAEALQELGLPVDPTAPLFAFGRLEBEQKGVLDIILAL 346
 DB 529 YLTSFGTYNLYNLTQTKRQCKAALQRELGLPVREDVPLIISFIRGLDHQKGVLDIAEAI 588
 QY 347 PKIATPKVQIATLGTGAAYEKLYNAIGTKYKRAKGVKVFSAFLAHMLTAGADFMVLP 406
 DB 589 PWM--SHDVQLVLMGTGRADLEQMLKEAFEQCHDKIRSWGVSVMARHITAGSDILLMP 647
 QY 407 SRFEPCGLTQLHAMHYGTPVVPVASTGGLVDTV 438
 DB 648 SRFEPCGLNQLYAMSYGIVPVVHVGGLRDTV 679

Search completed: June 4, 2003, 14:53:22
 Job time : 10.8812 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 14:35:26 ; Search time 21.7603 Seconds
(without alignments)
4147.394 Million cell updates/sec

Title: US-09-980-771A-7
Perfect score: 2288
Sequence: 1 ALDVMVAEVAPEWKTGGL.....AMHYGTVFVASTGGLVDTV 438

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2284	99.8	708	10 064925	064925 chlamydomon
2	1282	56.0	613	10 093YB1	093YB1 pisum sativ
3	1275	55.7	604	10 09FR03	09fr03 perilla fru
4	1261	55.1	608	10 093VD9	093vd9 ipomoea bat
5	1254	54.8	607	10 092S05	092sq5 astragalus
6	1250	54.6	606	10 09XIS6	09xis6 phaseolus v
7	1242	54.3	610	10 09MAQ0	09maq0 arabidopsis
8	1232.5	53.9	603	10 08SA49	08sa49 hordeum vul
9	1229	53.7	609	10 094LY7	094ly7 oryza sativ
10	1221.5	53.4	605	10 09SKK3	09skk3 triticum ae
11	1221.5	53.4	605	10 09FUU6	09fuu6 triticum ae
12	1221	53.4	609	10 08S9C4	08s9c4 oryza sativ
13	1217.5	53.2	604	10 09SLS6	09sls6 triticum tu
14	1216.5	53.2	574	10 09SYU0	09syu0 triticum ae
15	1216.5	53.2	604	10 09S7N5	09s7n5 triticum ae
16	1214.5	53.1	608	10 08YU1	08yu1 oryza sativ

17	1212.5	53.0	534	10 08W2G8	08w2g8 triticum ae
18	1211.5	53.0	599	10 09SQ58	09sq58 triticum ae
19	1210.5	52.9	606	10 043012	043012 oryza sativ
20	1209.5	52.9	604	10 09SL57	09sls7 triticum tu
21	1208.5	52.8	605	10 09SQ51	09sq51 aegilops sp
22	1207.5	52.8	605	10 09SL58	09sls8 triticum tu
23	1204.5	52.6	604	10 09SXX4	09sxx4 triticum ae
24	1202.5	52.6	604	10 09SL59	09sls9 triticum tu
25	1197.5	52.3	605	10 09SQ52	09sq52 triticum mo
26	1105.5	48.3	565	10 09XEN9	09xen9 triticum ae
27	895.5	39.1	313	10 09LKD3	09lkd3 vaquelinia
28	893.5	39.1	313	10 09LKE0	09lke0 prunus virg
29	893.5	39.1	313	10 09FYU8	09fyu8 aruncus dio
30	886.5	38.7	313	10 09LKE8	09lke8 kageneckia
31	885.5	38.7	313	10 09LKD1	09lkd1 exochorda r
32	879.5	38.4	313	10 09LKF5	09lkf5 amelanchier
33	879.5	38.4	313	10 09LKD6	09lkd6 vaquelinia
34	878.5	38.4	313	10 09LKE7	09lke7 kageneckia
35	878.5	38.4	313	10 09LKE1	09lke1 prinsepia s
36	877.5	38.4	313	10 09LKB6	09lkb6 malus sarge
37	877.5	38.4	313	10 09LKB8	09lkb8 rosa multif
38	876.5	38.3	313	10 09LKE5	09lke5 oemleria ce
39	876.5	38.3	313	10 09LKE4	09lke4 osteomeles
40	875.5	38.3	313	10 09LKF2	09lkf2 chaenomeles
41	874.5	38.2	313	10 09LKF3	09lkf3 amelanchier
42	874.5	38.2	313	10 09LKE2	09lke2 physocarpus
43	873.5	38.2	313	10 09LKF6	09lkf6 amelanchier
44	870.5	38.0	313	10 09LKD4	09lkd4 vaquelinia
45	870	38.0	623	10 09SSY9	09ssy9 oryza sativ

ALIGNMENTS

RESULT 1

064925
ID 064925 PRELIMINARY: PRT; 708 AA.
AC 064925;

DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Granule-bound starch synthase I precursor (bc 2.4.1.21).
GN STA2.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_taxID=3055;

RA Ball S.G.;
RA "Cloning of a cDNA encoding for the GBSSI in the green alga
Chlamydomonas reinhardtii.";
Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.
RC STRAIN=137C;
RA D'Hulst C., Watebled F., Ral J.-P., Abel G.J., Kossmann J.,
Ball S.G.;

RT "Cloning of a cDNA encoding for the GBSSI in the green alga
Chlamydomonas reinhardtii.";
Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [2]

RP SEQUENCE FROM N.A.
RA Watebled F., Ball S.G., D'Hulst C.;
RA "Granule-bound starch synthase I: A major enzyme involved in the
biogenesis of B-crystallites in starch granules.";
Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF026420; AAC17969.3; -
DR EMBL; AF433156; AAL28128.1; -
DR InterPro; IPR001296; Glycosyltransf_1.
DR InterPro; IPR002114; HPR_SerP_site.
DR Pfam; PF00534; Glycosyltransf_1; 1.
DR PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.
DR Glycosyltransferase; Transferase; Transit peptide.
FT TRANSIT 1 57 POTENTIAL.
FT CHAIN 58 708 GRANULE-BOUND STARCH SYNTHASE I.
SQ SEQUENCE 708 AA; 74623 MW; 7D2A5A07D8606469 CRC64;

Query Match

99.8%; Score 2284; DB 10; Length 708;


```

Db 421 PKFI-DQNVQIILVGTGKIMEKQIQELEVTPGKAIGVAKFNSPLAHKIAGADFIVP 479
QY 407 SRPEPCGLIQLHAMHYGTVPVASTGGLVDIV 438
Db 480 SRPEPCGLVOLHAMPYGTVPVASTGGLVDIV 511

RESULT 6
Q9XIS6 PRELIMINARY; PRT; 606 AA.
AC Q9XIS6;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Granule-bound starch synthase I precursor (EC 2.4.1.21).
GN GBSSI.
OS Phaseolus vulgaris (Kidney bean) (French bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
OX NCBI_TaxID=3885;
RN [1]
SEQUENCE FROM N.A.
RA Isono N., Nozaki K., Ito H., Matsui H., Honma M.;
RA "Phaseolus vulgaris L. mRNA for Granule-Bound Starch Synthase I
RT (GBSSI).";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB029346; BA82346.1; -.
DR InterPro: IPR001296; Glycos_transf_1.
DR Pfam: PF00534; Glycos_transf_1; 1.
KW Glycosyltransferase; Transferase; Transit peptide.
FT TRANSIT 1 POTENTIAL.
SQ SEQUENCE 606 AA; 67080 MW; 346943F0F6DE4ADB CRC64;

Query Match 54.68; Score 1250; DB 10; Length 606;
Best Local Similarity 54.4; Pred. No. 3.8e-87;
Matches 246; Conservative 65; Mismatches 105; Indels 36; Gaps 8;

QY 2 LDIVMAAEVAPWSKTGGLGVDTGGLPIELVKRGHRVMTIAPRDYQYADADMTSVVDIM 61
Db 80 MNLIFVGAEVAPWSKTGGLGVDTGGLPSALAEHGRVMTSPRDYQYKDAWDTNVEVK 139
QY 62 G----EKVRFHSTKKGVHRVMDHPFLAKVWGKTGSKLYGPRSGADYLDNHRKRALFC 117
Db 140 VADRIETVRFHSTKKGVHRVMDHPFLAKVWGKTGSKLYGPRSGADYLDNHRKRALFC 199
QY 118 KAATEAARVLPF-----GP-GEDCVFVANDVHSAVLPVLLKDEYQPKGQFTKAKSVLA 169
Db 200 QAALFAPVNLNSNKFSGPYGEDVIFVANDVHSAVLPVLLKDEYQPKGQFTKAKSVLA 259
QY 170 IHNIAFGQRMEEAFKDTKLP---QAADFKLAFSDGYAKVYTEATPMEDEKPLTGKTY 226
Db 260 IHNISYGRHFEDFPLNLPNEYSADF---FTDGLK-----PVRG--- 299
QY 227 KKNLWKGIIAADKLVTVSPNYATEIAADAAGVVELDTVIRAKGIEGVNGMDIEEWP 286
Db 300 KKNLWKGIIAADKLVTVSPNYATEIAADAAGVVELDTVIRAKGIEGVNGMDIEEWP 359
QY 287 KTDKFLSAPYQNSVYAGKAAKAEALQELGIPVDPPTAFPLAFIGRLEEQKGVDIILAAAL 346
Db 360 KTDKFDIHEDTTSVKRAKFLKALQAEVLGPNRDIPLIGFIRLEEQKGVDIILVEAI 419
QY 347 PKILATPKVQIATLGTKAAEKLVAIGTKYGRKAVKVFSAFLAHMLTAGADFMVLP 406
Db 420 PKFI-DQNVQIILVGTGKIMEKQIQELEVTPGKAIGVAKFNSPLAHKIAGADFIVP 478
QY 407 SRPEPCGLIQLHAMHYGTVPVASTGGLVDIV 438
Db 479 SRPEPCGLVOLHAMPYGTVPVASTGGLVDIV 510

RESULT 7

```

```

Q9MAQO PRELIMINARY; PRT; 610 AA.
AC Q9MAQO;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Granule-bound starch synthase.
GN F9L11.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Nguyen M., Lam B., Southwick A., Bei, Buehler E., Chin C.,
RA Chlou J., Choi E., Dunn P., Gonzalez A., Howng B., Kim C., Koo T.,
RA Lee J.M., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsy N.,
RA Pham P., Sakano H., Schwartz J., Shinn P., Thaveri A., Toriumi M.,
RA Vaysberg M., Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC006424; AAF31273.1; -.
DR InterPro: IPR001296; Glycos_transf_1.
DR Pfam: PF00534; Glycos_transf_1; 1.
SQ SEQUENCE 610 AA; 66879 MW; CF17F25BBL2220DF CRC64;

Query Match 54.3%; Score 1242; DB 10; Length 610;
Best Local Similarity 55.3%; Pred. No. 1.6e-86;
Matches 250; Conservative 59; Mismatches 107; Indels 36; Gaps 9;

QY 2 LDIVMAAEVAPWSKTGGLGVDTGGLPIELVKRGHRVMTIAPRDYQYADADMTSVVDI- 60
Db 84 MSVIFGAEVGWSKTGGLGVDTGGLPPALAEHGRVMTICPRDYQYKDAWDTCVVVOIK 143
QY 61 MGEK--VRYFHSSTKKGVHRVMDHPFLAKVWGKTGSKLYGPRSGADYLDNHRKRALFC 117
Db 144 VGDKVENVRFFHSTKKGVHRVMDHPFLAKVWGKTGSKLYGPRSGADYLDNHRKRALFC 203
QY 118 KAATEAARVLPF-----GP-GEDCVFVANDVHSAVLPVLLKDEYQPKGQFTKAKSVLA 169
Db 204 QAALFAPVNLNSNKFSGPYGEDVIFVANDVHSAVLPVLLKDEYQPKGQFTKAKSVLA 263
QY 170 IHNIAFGQRMEEAFKDTKLP---QAADFKLAFSDGYAKVYTEATPMEDEKPLTGKTY 226
Db 264 IHNIAFGQRMEEAFKDTKLP---QAADFKLAFSDGYAKVYTEATPMEDEKPLTGKTY 303
QY 227 KKNLWKGIIAADKLVTVSPNYATEIAADAAGVVELDTVIRAKGIEGVNGMDIEEWP 286
Db 304 KKNLWKGIIAADKLVTVSPNYATEIAADAAGVVELDTVIRAKGIEGVNGMDIEEWP 363
QY 287 KTDKFLSAPYQNSVYAGKAAKAEALQELGIPVDPPTAFPLAFIGRLEEQKGVDIILAAAL 346
Db 364 STDYIDIKYDITVTDAKPLIKEALQAEVLGPNRDIPLIGFIRLEEQKGVDIILVEAI 423
QY 347 PKILATPKVQIATLGTKAAEKLVAIGTKYGRKAVKVFSAFLAHMLTAGADFMVLP 406
Db 424 SKFMGL-NVQVILGTGKIMEKQIQELEVTPGKAIGVAKFNSPLAHMLTAGADFMVLP 482
QY 407 SRPEPCGLIQLHAMHYGTVPVASTGGLVDIV 438
Db 483 SRPEPCGLIQLHAMHYGTVPVASTGGLVDIV 514

RESULT 8
Q8SA49 PRELIMINARY; PRT; 603 AA.
AC Q8SA49;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Granule-bound starch synthase.

```

```

GN 259116.5.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. MEREX;
RA Ma J., Samiguel P.J., Dubcovsky J., Shiloff B.A., Rostoks N.,
RA Bennettzen J.L.;
RA Jiang Z., Busso C.S., Kleinohs A., Devos K.M., Ramakrishna W.,
RA *Comparative sequence analysis of wx1 homologous regions in barley,
RT maize, pearl millet, rice, sorghum and diploid wheat.*;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF474373; AAL77109.1; -.
SQ SEQUENCE 603 AA; 66280 MW; 467A3D56A82125CB CRC64;

Query Match 53.9%; Score 1232.5; DB 10; Length 603;
Best Local Similarity 54.1%; Pred. No. 8.1e-86;
Matches 243; Conservative 60; Mismatches 117; Indels 29; Gaps 6;

QY 2 LDIVVAAEVAPWSKTGGLDVTGGGLPIELVKRHRVMTIAPRYDQYADAWDTSVVVDIM 61
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
76 MNLVFGAEMAPWSKTGGLDVTGGGLPPMAAANGHRVMVSPRYDQYKDAWTSVISEIK 135
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
62 ----GEKVRYPHSIKKGVRHVWIDHPWFLAKVWGKTGSKLYGPRGADYLDNHRKRALFC 117
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
136 VADEYRVERVFFHCYKRGVDRVFDHPWFLAKVWGKTGSKLYGPRGADYLDNHRKRALFC 195
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
118 KAAIEARVL-----PFGP---GECVFFVANDWHSALVPLVKDEYQPKQFTKAKSVLA 169
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
196 QAALEAPRLINNNPYFSGPVEDVFCNDWHTGLLACVLSNQSGIYRTAKVAPC 255
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
170 IHNIAFGQRMWEEAFKDKLQAAEDKLAFSQYAKVYTEATPMEDEKPPLTGKTKYKI 229
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
256 IHNISYQGRFDFDQAQLNLPDRFKSSFDIDGYDK-----PVEG---RKI 298
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
230 NMLKGGIITADKLVTVSPNYATEIAADAAGGVLDVTIRAKGIEGIVNGMDIEWNPKT 289
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
299 NWMKAGILQADKLVTVSPNYATEIAADAAGGVLDVTIRAKGIEGIVNGMDIEWNPKT 358
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
290 KFLSAPYDQNSVYAGKAAKALQAEGLPVDPTAPLFAFGRLERQKGVDTIILAALPKI 349
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
359 KFLAVNYDITTALEAKALNKEALQAEGLPVDKVPVYAFVGRLEEQKGPDMVAAIPEI 418
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
350 LATPKVQITAILGTGKAAEKLVAIGTKYKRAKGVVKSAPLAHMLTAGADFMVPSRF 409
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
419 LKEEDVQIILGTGKKKFKLLKSMEEKFPGKVRVAVRFNAPLAHOMMAGADLLAVTSRF 478
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
410 EPCGLIQLHAMHYGTPVAVSTGGGLVDTV 438
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
479 EPCGLIQLGMRGTPVCACSTGGGLVDTV 507
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3

RESULT 9
Q94LY7
ID Q94LY7 PRELIMINARY; PRT; 609 AA.
AC Q94LY7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Granule-bound starch synthase (EC 2.4.1.21).
GN WAXY.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. REXMONT;
RA Larkin P.D., McClung A.M., Ayres N.M., Park W.D.;

```

```

RT "The Wx locus (Granule Bound Starch Synthase) is strongly associated
RT with pasting curve characteristics in rice (Oryza sativa L.).";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF141954; AAF72561.1; -.
DR InterPro; IPR001296; Glycos_transf_1.
DR InterPro; IPR000794; Glycos_transf_1; 1.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
SQ SEQUENCE 609 AA; 66466 MW; 6D2615EB203CBDC8 CRC64;

Query Match 53.7%; Score 1229; DB 10; Length 609;
Best Local Similarity 53.9%; Pred. No. 1.5e-85;
Matches 242; Conservative 61; Mismatches 116; Indels 30; Gaps 6;

QY 2 LDIVVAAEVAPWSKTGGLDVTGGGLPIELVKRHRVMTIAPRYDQYADAWDTSVVVDI- 60
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
83 MNLVFGAEMAPWSKTGGLDVTGGGLPPMAAANGHRVMVSPRYDQYKDAWTSVVAEIK 142
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
61 ----MGKVRYPHSIKKGVRHVWIDHPWFLAKVWGKTGSKLYGPRGADYLDNHRKRALFC 117
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
143 VADRYRVERVFFHCYKRGVDRVFDHPWFLAKVWGKTGSKLYGPRGADYLDNHRKRALFC 202
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
118 KAAIEARVL-----PFGP---GECVFFVANDWHSALVPLVKDEYQPKQFTKAKSVLA 169
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
203 QAALEAPRLINNNPYFSGPVEDVFCNDWHTGLLACVLSNQSGIYRTAKVAPC 262
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
170 IHNIAFGQRMWEEAFKDKLQAAEDKLAFSQYAKVYTEATPMEDEKPPLTGKTKYKI 229
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
263 IHNISYQGRFDFDQAQLNLPDRFKSSFDIDGY-----DTPVEG-----RKI 305
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
230 NMLKGGIITADKLVTVSPNYATEIAADAAGGVLDVTIRAKGIEGIVNGMDIEWNPKT 289
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
306 NWMKAGILEADKLVTVSPNYATEIAADAAGGVLDVTIRAKGIEGIVNGMDIEWNPKT 365
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
290 KFLSAPYDQNSVYAGKAAKALQAEGLPVDPTAPLFAFGRLERQKGVDTIILAALPKI 349
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
366 KIITAKYDITTALEAKALNKEALQAEGLPVDKVPVYAFVGRLEEQKGPDMVAAIPEL 425
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
350 LATPKVQITAILGTGKAAEKLVAIGTKYKRAKGVVKSAPLAHMLTAGADFMVPSRF 409
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
426 M-QEDVQIILGTGKKKFKLLKSMEEKYKVRVAVRFNAPLAHMLTAGADFMVPSRF 484
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
410 EPCGLIQLHAMHYGTPVAVSTGGGLVDTV 438
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3
485 EPCGLIQLGMRGTPVCACSTGGGLVDTV 513
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 3

RESULT 10
Q9SXX3
ID Q9SXX3 PRELIMINARY; PRT; 605 AA.
AC Q9SXX3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Starch synthase (GBSSI) (EC 2.4.1.21).
GN WAXY.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99321800; PubMed=10393240;
RT Murai J., Taira T., Ohta D.;
RT "Isolation and characterization of the three Waxy genes encoding the
RT granule-bound starch synthase in hexaploid wheat.";
RL Gene 234:71-79(1999).
DR EMBL; AB019623; BAA77351.1; -.
DR InterPro; IPR001296; Glycos_transf_1.
DR InterPro; IPR000794; Ketoacyl-synt.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.

```



```

QY 170 IHNIAFOGRMWEAFKDTKLPQAAFDKLAISDGYAKVYTEATPMEDEKPLTGKTYKKI 229
Db 263 IHNISYOGREFADYPELNSERFSRSDFDIDGY-----DTPVEG-----RKI 305

QY 230 NWLKGGLIADKLVTVSPNYATEIAADAAGGVLDTVIRAKIGIEGVNGMDIEWNPDKTD 289
Db 306 NWMKAGILEADRVLTSPYAEELISGARCELDNIMRLTGITGVNGMDVSEMDPSK 365

QY 290 KFLSAPYDQNSVYAGKAAAEALQAEGLPVDPTAPLFAFVIRLEEGKGVDDIILAALPKI 349
Db 366 KYITAKYDATTAEAKALNEALQAEAGLPVDRKIPLIAFIATIGRLEEGKGVDDVMAAIP 425

QY 350 LATEPKVOITAILGTGKAAEKLVAIKYKGRAGVVKFSPAPLAHMLTAGADFMVPSRF 409
Db 426 M-QEDVQIVLGTGKKFELLSKVEEKYPGKRVAVVFNAPLAHMLTAGADVLAAPS 484

QY 410 EPGGLIQLHAMHYGTPVWVASTGGGLVDTV 438
Db 485 EPGGLIQLGMRGTPCACASTGGGLVDTV 513

RESULT 13
Q9SL56 PRELIMINARY; PRT; 604 AA.
AC Q9SL56;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Starch synthase (GBSSI) (EC 2.4.1.21).
GN WAXY.
OS Triticum turgidum subsp. durum (durum wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4567;
RN [1]
RP SEQUENCE FROM N.A.
RA Murai J., Taira T., Ohta D.;
RT "Isolation and characterization of the four Waxy genes encoding the
granule-bound starch synthase in tetraploid wheats.";
RL Appl. Biol. Sci. 0:0-0(1999).
DR EMBL; AB029064; BAA88512.1; -.
DR InterPro; IPR001296; Glycosyltransf_1.
DR InterPro; IPR000794; Ketoacyl-synt.
DR Pfam; PF00534; Glycosyltransf_1; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 604 AA; 66227 MW; 3509B1E28D9D9E45 CRC64;

```

```

Query Match 53.2%; Score 1217.5; DB 10; Length 604;
Best Local Similarity 53.9%; Pred. No. 1.1e-84;
Matches 242; Conservative 58; Mismatches 120; Indels 29; Gaps 6;

QY 2 LDIVMAAEVAPWSKTGGLGVDVGTGLPIELVKRHRVMTIAPRYDQYADADTSSVVDIM 61
Db 77 MNLVFGAEMAPWSKTGGLGVDVGLGLPPAMAANGHRVMTISPRDQYKDADTSSVSEIK 136

QY 62 ----GKVRVPHSTIKGVHVRWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKRALFC 117
Db 137 VADEYERVRVPHCYKRGVDRVFDHPGCFLEKVRGKTEKIYGPAGTDYEDNQLRFSLLC 196

QY 118 KAATEAARVL-----PF--GP-GEDCVFVANDHNSALVPVLLKDEYQPKQFTKAKSVLA 169
Db 197 QAALAEPRILDLNNPFGPYGDEVVFCNDWHTGLLACYLKSNSYGSNGIYTAKVAF 256

QY 170 IHNIAFOGRMWEAFKDTKLPQAAFDKLAISDGYAKVYTEATPMEDEKPLTGKTYKKI 229
Db 257 IHNISYOGREFADYPELNSERFSRSDFDIDGYDK-----PVEG-----RKI 299

QY 230 NWLKGGLIADKLVTVSPNYATEIAADAAGGVLDTVIRAKIGIEGVNGMDIEWNPDKTD 289
Db 300 NWMKAGILQADKLVLTSPYAEELISGARCELDNIMRLTGITGVNGMDVSEMDPTKD 359

QY 290 KFLSAPYDQNSVYAGKAAAEALQAEGLPVDPTAPLFAFVIRLEEGKGVDDIILAALPKI 349
Db 366 KYITAKYDATTAEAKALNEALQAEAGLPVDRKIPLIAFIATIGRLEEGKGVDDVMAAIP 425

QY 350 LATEPKVOITAILGTGKAAEKLVAIKYKGRAGVVKFSPAPLAHMLTAGADFMVPSRF 409
Db 426 M-QEDVQIVLGTGKKFELLSKVEEKYPGKRVAVVFNAPLAHMLTAGADVLAAPS 484

```

```

QY 290 KFLSAPYDQNSVYAGKAAAEALQAEGLPVDPTAPLFAFVIRLEEGKGVDDIILAALPKI 349
Db 360 KFLAVNYDITTALEGKALNEALQAEVGLPVDKRVPLVAFVIRLEEGKGVDDVMAAIP 419

QY 350 LATEPKVOITAILGTGKAAEKLVAIKYKGRAGVVKFSPAPLAHMLTAGADFMVPSRF 409
Db 420 LKEEDVQIVLGTGKKFELLSKVEEKYPGKRVAVVFNAPLAHMLTAGADVLAAPS 479

QY 410 EPGGLIQLHAMHYGTPVWVASTGGGLVDTV 438
Db 480 EPGGLIQLGMRGTPCACASTGGGLVDTV 508

RESULT 14
Q9SYU0 PRELIMINARY; PRT; 574 AA.
AC Q9SYU0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Granule-bound starch synthase.
GN WX-D1.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. WAXY;
RX MEDLINE=99254805; PubMed=10323226;
RA Vrinten P., Nakamura T., Yamamori M.;
RT "Molecular characterization of waxy mutations in wheat.";
RL Mol. Gen. Genet. 261:463-471(1999).
DR EMBL; AF113844; AAD26156.1; -.
DR InterPro; IPR001296; Glycosyltransf_1.
DR InterPro; IPR000794; Ketoacyl-synt.
DR Pfam; PF00534; Glycosyltransf_1; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
SQ SEQUENCE 574 AA; 63267 MW; 87C13579508CB402 CRC64;

```

```

Query Match 53.2%; Score 1216.5; DB 10; Length 574;
Best Local Similarity 53.9%; Pred. No. 1.3e-84;
Matches 242; Conservative 58; Mismatches 120; Indels 29; Gaps 6;

QY 2 LDIVMAAEVAPWSKTGGLGVDVGTGLPIELVKRHRVMTIAPRYDQYADADTSSVVDI 60
Db 77 MNLVFGAEMAPWSKTGGLGVDVGLGLPPAMAANGHRVMTISPRDQYKDADTSSVSEIK 136

QY 61 ---MGEKVRVPHSTIKGVHVRWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKRALFC 117
Db 137 VVDYERVRVPHCYKRGVDRVFDHPGCFLEKVRGKTEKIYGPAGTDYEDNQLRFSLLC 196

QY 118 KAATEAARVL-----PF--GP-GEDCVFVANDHNSALVPVLLKDEYQPKQFTKAKSVLA 169
Db 197 QAALAEPRILDLNNPFGPYGDEVVFCNDWHTGLLACYLKSNSYGSNGIYTAKVAF 256

QY 170 IHNIAFOGRMWEAFKDTKLPQAAFDKLAISDGYAKVYTEATPMEDEKPLTGKTYKKI 229
Db 257 IHNISYOGREFADYPELNSERFSRSDFDIDGYDK-----PVEG-----RKI 299

QY 230 NWLKGGLIADKLVTVSPNYATEIAADAAGGVLDTVIRAKIGIEGVNGMDIEWNPDKTD 289
Db 300 NWMKAGILQADKLVLTSPYAEELISGARCELDNIMRLTGITGVNGMDVSEMDPTKD 359

QY 290 KFLSAPYDQNSVYAGKAAAEALQAEGLPVDPTAPLFAFVIRLEEGKGVDDIILAALPKI 349
Db 360 KFLAVNYDITTALEGKALNEALQAEVGLPVDKRVPLVAFVIRLEEGKGVDDVMAAIP 419

QY 350 LATEPKVOITAILGTGKAAEKLVAIKYKGRAGVVKFSPAPLAHMLTAGADFMVPSRF 409
Db 420 LKEEDVQIVLGTGKKFELLSKVEEKYPGKRVAVVFNAPLAHMLTAGADVLAAPS 479

```

```
QY 410 EPCGLIOLHAMHYGTVPVASTGGLVDTV 438
||||| | | | | | | | | | | | | | | | |
Db 480 EPCGLIOLGMRVGTTPCACASTGGLVDTI 508

RESULT 15
Q9S7N5
ID 09S7N5 PRELIMINARY; PRT; 604 AA.
AC 09S7N5; 2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Starch synthase (GBSSI) (EC 2.4.1.21) (Granule-bound starch synthase
I).
DE WAXY OR WX-TD OR WX-DI.
GN Triticum aestivum (Wheat), and
OS Triticum aestivum (Wheat), and
OS Aegilops tauschii (Aegilops squarrosa).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565, 37682;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=T.aestivum;
RX MEDLINE=93321800; PubMed=10393240;
RA Murai J., Taira T., Ohta D.;
RT "Isolation and characterization of the three waxy genes encoding the
granule-bound starch synthase in hexaploid wheat.";
RL Gene 234:71-79(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=A.squarrosa; STRAIN=CV. CPI 110799;
RA Yan L., Bhawe M., Fairclough R., Konic C., Rahman S., Appels R.;
RT "The genes encoding granule-bound starch synthases at the waxy loci of
the A, B and D progenitors of common wheat.";
RL Genome 0:0-0(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=T.aestivum; STRAIN=CV. CHINESE SPRING;
RX MEDLINE=20098733; PubMed=10631269;
RA Vrinten P.L., Nakamura T.;
RT "Wheat granule-bound starch synthase I and II are encoded by separate
genes that are expressed in different tissues.";
RL Plant Physiol. 122:255-264(2000).
DR EMBL; AB019624; BAA7352.1; -
DR EMBL; AF110375; AAF06938.1; -
DR EMBL; AF163319; AAF34135.1; -
DR InterPro; IPR001296; Glycos_transf_1.
DR InterPro; IPR000794; Ketoacyl-synt.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR PROSITE; PS00606; B.KETOACYL_SYNTHASE; UNKNOWN_1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 604 AA; 66269 MW; 6F69B3AB6594AE49 CRC64;

Query Match 53.2%; Score 1216.5; DB 10; Length 604;
Best Local Similarity 53.9%; Pred. No. 1.4e-84;
Matches 242; Conservative 58; Mismatches 120; Indels 29; Gaps 6;

QY 2 LDIVVAAEAPWSKGGTGGTGLPLELVKGRHYMTIAPRYDQYADAMDTSVVVDI- 60
:::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 77 MNLVFGAEAPWSKGGTGGTGGTGLPLELVKGRHYMTIAPRYDQYADAMDTSVVSEIK 136
:::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 ---MGEKRVFHSIKKGVRHWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKFFALFC 117
:::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 137 VVDKYERVRYFHCYKRGVDRVFDHPCFLEKVRCKTEKIYGPDA GTDYEDNQORFSLC 196
:::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 118 KAAIEARVL-----PF--GP-GEDEVANDHSAVPVLLKDEYQPKGQFTAKSVIA 169
:::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 197 QAALEVPRIILNDNNPFGSGYGEDVVFVNCNDHTGLIACYLKSNGYIYRAAKVAF 256
:::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 170 IHNIAFGRMWEAEAFKDKLPQAAFDKLA FSDGYAKVYTEATPMEDEKPPLTGKTKYKI 229
||||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
Db 257 IHNISYGRFSFDDPFAQLNLPDRFKSSDFDIDGYDK-----PVEG---RKI 299
QY 230 NWLKGGIIAADKLVTVSPNYATEIAADAAGVGVELDTVIRAKGIEGIVNGMDIEWNPKTD 289
||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 300 NWMKAGILQADKLVTVSPNYATEIAADAAGVGVELDTVIRAKGIEGIVNGMDIEWNPKTD 359
||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 290 KFLSAPYDQNSVYAGKAAAKEALQAEGLPVDPTAPLFAFGRLEEQKGVDIILAAALPKI 349
||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 360 KFLAVNYDITTALEGGKALNKEALQAEGLPVDKVPVLFVAFGRLEEQKGVDMIAAIPDI 419
||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 350 LATPKVQITAILGTGKAAAYEKLVAIGTKYKRAKGVYKFSAPLAHMLTAGADFMVPSRF 409
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 420 LKEEDVQIVLLGTGKKKFERLLKSTEEKFPSKRVAVRNFNAPLAHQMMAGADVLAITSRF 479
||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 410 EPCGLIOLHAMHYGTVPVASTGGLVDTV 438
||||| | | | | | | | | | | | | | | | |
Db 480 EPCGLIOLGMRVGTTPCACASTGGLVDTI 508
||||| | | | | | | | | | | | | | | | |

Search completed: June 4, 2003, 14:56:06
Job time : 26.7603 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 14:49:26 ; Search time 7.79255 Seconds
(without alignments)
1653.790 Million cell updates/sec

Title: US-09-980-771A-7
Perfect score: 2288
Sequence: 1 ALDVMVAEAPWSKTGGL.....AMHYGTVPVVASTGLYDVT 438

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/6C.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1229	53.7	609	3	US-08-941-445A-7
2	1188	51.9	600	4	US-09-388-743-22
3	1185	51.8	533	3	US-08-941-445A-5
4	1164	50.9	616	4	US-09-388-743-14
5	1154	50.4	614	4	US-09-388-743-18
6	1147	50.1	615	4	US-09-388-743-2
7	870.5	38.0	641	4	US-08-836-567-10
8	862	37.7	671	4	US-09-196-390-2
9	861	37.6	649	4	US-09-192-909-2
10	859	37.5	539	3	US-08-941-445A-21
11	859	37.5	583	3	US-08-941-445A-13
12	833	36.4	801	4	US-09-388-743-26
13	826	36.1	690	4	US-09-388-743-6
14	826	36.1	767	4	US-08-836-567-8
15	817	35.7	799	4	US-09-196-390-6
16	810	35.4	698	3	US-08-941-445A-11
17	800	35.0	558	4	US-08-836-567-6
18	696	30.4	459	4	US-08-836-567-4
19	671.5	29.3	669	3	US-08-941-445A-9
20	558	24.4	477	1	US-07-735-065-2
21	558	24.4	477	1	US-08-469-202-12
22	558	24.4	477	2	US-08-484-434C-12
23	528.5	23.1	735	4	US-09-115-704-2
24	528	23.1	677	4	US-08-836-567-2
25	528	23.1	1197	4	US-08-836-567-12
26	528	23.1	1230	2	US-08-968-542C-35
27	499	21.8	533	4	US-09-388-743-10

28	471.5	20.6	1674	2	US-08-968-542C-12	Sequence 12, Appl
29	240	10.5	64	2	US-08-470-720-15	Sequence 15, Appl
30	215.5	9.4	79	2	US-08-470-720-13	Sequence 13, Appl
31	143.5	6.3	59	2	US-08-470-720-14	Sequence 14, Appl
32	129	5.6	30	2	US-08-470-720-9	Sequence 9, Appl
33	121	5.3	27	2	US-08-470-720-7	Sequence 7, Appl
34	112	4.9	34	2	US-08-470-720-11	Sequence 11, Appl
35	110.5	4.8	1037	4	US-09-134-001C-4794	Sequence 6, Appl
36	107	4.7	69	2	US-08-470-720-6	Sequence 18, Appl
37	107	4.7	111	2	US-08-470-720-18	Sequence 41, Appl
38	105.5	4.6	490	4	US-09-292-225-41	Sequence 35, Appl
39	105.5	4.6	509	4	US-09-292-225-35	Sequence 38, Appl
40	105.5	4.6	509	4	US-09-292-225-38	Sequence 1, Appl
41	100.5	4.4	387	4	US-09-457-302-1	Sequence 5024, Ap
42	99	4.3	494	2	US-09-134-001C-5024	Sequence 14, Appl
43	98	4.3	4545	2	US-08-804-227C-14	Sequence 8, Appl
44	98	4.3	4550	2	US-08-804-227C-8	Sequence 2, Appl
45	98	4.3	4550	2	US-08-804-198-2	

ALIGNMENTS

RESULT 1
US-08-941-445A-7
; Sequence 7, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-941-445A-7

Query Match 53.7%; Score 1229; DB 3; Length 609;
Best Local Similarity 53.7%; Pred. No. 3,7e-120;
Matches 242; Conservative 61; Mismatches 116; Indels 30; Gaps 6;
QY 2 LDIVMVAEAPWSKTGGLDVTGGLPIELVKCRHRYMTAPRYDQYADAWDTSVVYDI- 60
Db 83 MNVVFGEAPWSKTGGLDVLGLGPPMAAANGHRYVWVLSPRYDQYKDAWDTSWAEIK 142


```

QY 170 IHNTAFQGRMWEAFKDTKLQAAFDKLAFLSDGAKYVTEATPMEDEKPLTGKTYKKI 229
DB 185 IHNTSYQGRFAFSYDPELNLPERFKSSDFDIDGEEK-----PVGE---RKI 227

QY 230 NNLXGGIIAADKLVTVSPNYATEIAADAAGGVELDTVIRAKGIEGIVNGMDIEWNPKTD 289
DB 228 NNMKAGILEADRVLTSPYAAELISGTARGCENLNMRLGITGIVNGMDVSEWDPSPRD 287

QY 290 KFLSAPYDONSYYAGKAAKALQALGELPLVDPTAPLFAFAGRLGEEQKGVDIILAAALPKI 349
DB 288 KYIAVKYDVSTAVEAKALNKEALQAEVGLPVDNRNPLVAFIAGRLGEEQKGVDMVMAAIPQL 347

QY 350 L-ATPKVQIAILGTCKAAEYKLVNAIGTKYKGRAGVVKFSAPLAHMLTAGADMLVPSRF 408
DB 348 MEMVEDQIVLIGTGKKFERMLMSAEKFFPKGRVAVVYKFNAAALAHMAGADVLAVTSR 407

QY 409 EPCGLIQLHAMHYGTVPVVASTGGLVDTV 438
DB 408 EPCGLIQLQGMRYGTCPACASTGGLVDTI 437

RESULT 4
US-09-388-743-14
; Sequence 14, Application US/09388743
; Patent No. 6423886
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Ian
; TITLE OF INVENTION: No. 6423886el Starch Synthese Polynucleotides and Their
; FILE OF INVENTION: Use in the Production of New Starches
; CURRENT APPLICATION NUMBER: US/09/388,743
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Canna edulis
US-09-388-743-14

Query Match 50.9%; Score 1164; DB 4; Length 616;
Best Local Similarity 51.9%; Pred. No. 2.5e-113;
Matches 233; Conservative 60; Mismatches 126; Indels 30; Gaps 7;

QY 2 LDIWVAAEVAPWSKTGGLGVDVTGGLPIELVKRHRVMTIAPRYDQYADAMDTSVVVDI- 60
DB 90 MNLVFGVGEVAPWSKTGGLGVDVLRGLPPAMAAIGHRVMTVVPVYDQYKDIWDTSPVVELK 149

QY 61 MGEK---VRYFHSITKGVHRVWIDHPWFLAKVWGTGSKLYGPRSGADYLDNHKRFALFC 117
DB 150 VGDKIEVRFHCYKRGVDRVFDHPFLEKRVWGTGSKLYGPGVTGDIADNQLRFSLLC 209

QY 118 KAAIEAARVLFF-----GP-GEDCVFVANDHSAVLPVLLKDEYQPKGQFTKAKSVLA 169
DB 210 LALEAPRLNLSNKKYISGPGYDGVFIANDHSAVLLPCVLKTYQSHGYMNAKVAF 269

QY 170 IHNTAFQGRMWEAFKDTKLQAAFDKLAFLSDGAKYVTEATPMEDEKPLTGKTYKKI 229
DB 270 IHNTAYQGRFAFSDFELNLPNFKKSFDFMDGYDK-----PVKG---RKI 312

QY 230 NNLXGGIIAADKLVTVSPNYATEIAADAAGGVELDTVIRAKGIEGIVNGMDIEWNPKTD 289
DB 313 NNMKAGIIECDRLTVSPYAAELISGTARGCENLNMRLGITGIVNGMDTTEWNPPLTD 372

QY 290 KFLSAPYDONSYYAGKAAKALQALGELPLVDPTAPLFAFAGRLGEEQKGVDIILAAALPKI 349
DB 373 KYISTNYDATVLDKAPLCREALQAEGLPVNKNKVLAVFVGRLEDEQKGSILAAAPL 432

QY 350 L-ATPKVQIAILGTCKAAEYKLVNAIGTKYKGRAGVVKFSAPLAHMLTAGADMLVPSRF 409
DB 433 LC-ENVQVIVLGTGKKLESELTLEEMFDPKFAHLKFNVPVLAHMAIMAGADILVPSRF 491

```

```

QY 410 EPCGLIQLHAMHYGTVPVVASTGGLVDTV 438
DB 492 EPCGLIQLQAMRYGTLPMCGSTTGGLVDTV 520

RESULT 5
US-09-388-743-18
; Sequence 18, Application US/09388743
; Patent No. 6423886
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Ian
; TITLE OF INVENTION: No. 6423886el Starch Synthese Polynucleotides and Their
; FILE OF INVENTION: Use in the Production of New Starches
; CURRENT APPLICATION NUMBER: US/09/388,743
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Typha latifolia
US-09-388-743-18

Query Match 50.4%; Score 1154; DB 4; Length 614;
Best Local Similarity 50.8%; Pred. No. 2.8e-112;
Matches 228; Conservative 62; Mismatches 129; Indels 30; Gaps 7;

QY 2 LDIWVAAEVAPWSKTGGLGVDVTGGLPIELVKRHRVMTIAPRYDQYADAMDTSVVVDI- 60
DB 88 MNLVFGVGAEMAPWSKTGGLGVDVLRGLPPALAAANGHRVVIAPRYDQYDAMDNDALVELK 147

QY 61 MG---EKVRYEHSIKKGVHRVWIDHPWFLAKVWGTGSKLYGPRSGADYLDNHKRFALFC 117
DB 148 VGDRCETVRFHCYKRGVDRVFDHPFLEKRVWGTGSKLYGPGVTGDIADNQLRFSLLC 207

QY 118 KAAIEAARVLFF-----GP-GEDCVFVANDHSAVLPVLLKDEYQPKGQFTKAKSVLA 169
DB 208 QAALEAPRLNLSNDSFSGPYGVEDVIFICNDWHSTLLPCVLKSMYHPRGIVKNAKVAF 267

QY 170 IHNTAFQGRMWEAFKDTKLQAAFDKLAFLSDGAKYVTEATPMEDEKPLTGKTYKKI 229
DB 268 IHNTSYQGRFAFSDFELNLPNFKKSFSDIDGYNK-----PVKG---MKI 310

QY 230 NNLXGGIIAADKLVTVSPNYATEIAADAAGGVELDTVIRAKGIEGIVNGMDIEWNPKTD 289
DB 311 NNMKAGIILESDRVFTVSPYAAELISGEERGVELDNLIRVTSITGIVNGMDVSEWNPPLTD 370

QY 290 KFLSAPYDONSYYAGKAAKALQALGELPLVDPTAPLFAFAGRLGEEQKGVDIILAAALPKI 349
DB 371 KYISVNYDAKPVMEAKPLNKEALQAESWLACRQGHPCNCHGRLEEQKGSILAAASIP 430

QY 350 L-ATPKVQIAILGTCKAAEYKLVNAIGTKYKGRAGVVKFSAPLAHMLTAGADMLVPSRF 409
DB 431 M-DENVQIILIGTGKKEMENQLESMEEMFPDKVAVMKNFAPLAHQMTAGADIIVPSRF 489

QY 410 EPCGLIQLHAMHYGTVPVVASTGGLVDTV 438
DB 490 EPCGLIQLQGMRYGTSPSACSSTTGGLVDTV 518

```

```

RESULT 6
US-09-388-743-2
; Sequence 2, Application US/09388743
; Patent No. 6423886
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Ian
; TITLE OF INVENTION: No. 6423886el Starch Synthese Polynucleotides and Their
; FILE OF INVENTION: Use in the Production of New Starches
; FILE REFERENCE: 1144
; CURRENT APPLICATION NUMBER: US/09/388,743

```

RESULT 7
US-08-836-567-10
; Sequence 10, Application US/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,567
; FILING DATE: 24-JUL-1997

```

RESULT 8
US-09-196-390-2
; Sequence 2, Application US/09196390
; Patent No. 6307125
; GENERAL INFORMATION:
; APPLICANT: Block, Martina
; APPLICANT: Lorz, Horst
; APPLICANT: Lutticke, Stephanie
; APPLICANT: Walter, Lennart
; APPLICANT: Froberg, Claus
; APPLICANT: Kossmann, Jens
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
; TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH
; TITLE OF INVENTION: SYNTHESIS

```



```

Db      362  NLKGAVVADRIYVTSKYSWEVTT-AEGGGGLNELLSSRKSVLNGVINGIDINDWNP 420
QY      288  TKFLSAPYDQNSVYAGKAAKEALQAEGLPVDPTAPLFAFGRLEEQKGVDIILAALP 347
Db      421  TDKICFCHSVDDL-SGKAKCKGALQKELGPIRPDVPILGFTIGRLDYQGDILQIILP 479
QY      348  KILATPKVQIAITLGTGKAAEKLVAIGTKYGRKGVVFKFSAPLAHMLTAGADFMVPS 407
Db      480  DLM-REDVQFVMLGSGDPELEDWMRSTESIFKDKRGWGFSPVSHRITAGCDILMPS 538
QY      408  REPCGLIQLHAMHYGTPVPVASTGGGLVDTV 438
Db      539  REPCGLNQLYAMQYGTVPVHATGGLRDTV 569

RESULT 10
US-08-941-445A-21
; Sequence 21, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanning
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 539 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-941-445A-21

```

```

Query Match      37.5%; Score 859; DB 3; Length 539;
Best Local Similarity 42.2%; Pred. No. 2.le-81;
Matches 190; Conservative 77; Mismatches 145; Indels 38; Gaps 12;

QY      4  IVMVAAEVPWSKGGGLGVDVTGGLPIELVKRGRHVMYTIAPRY-----DQYADAWDTSVV 57
Db      51  IVFTVGEASPYAKSGGLDVCGLPVALAARGHRVVMVMPRYLNGTSDKNYANAFYTEKH 110
QY      58  VDIM---GE-KVRFHSLKKGVRHVDIDHFWFLAKVWGKSGKLYGPRSGADYLDNHRF 113
Db      111  IRIFCFGGEHVEVTFEYRDSVDWVDFHFSY-----HRPGNLYGDKFGA-FGDNQFRY 163
QY      114  ALFCKAAEAAARVLPEFG---GEDCVFVANDWHSALVPVLLKDEYQPKGQFTKAKSVLAI 170

```

```

Db      164  TLICYAAACEAPLILELGGYIYGQCMFVNDWHDHSLVPLVLLAAKRYPGYVKDSRLVI 223
QY      171  HNIAFGQRMWEAEKFTKLPQAAFDKLAFLSDGVYAKVYTEATPMEDEKPLTIGTKYKIN 230
Db      224  HNLHGQGVFEPASTYDGLGLPEWYGALEW-----VPEWARRHALDKG-----EAVN 270
QY      231  WLKGGITIAADKLVTSVSNYATEIAADAAGGVELDTVI--RAKGIEGIVNGMDIEWNPKT 288
Db      271  FLKGAVVADRIYVTSKYSWEVTT-AEGGGGLNELLSSRKSVLNGVINGIDINDWNPAT 329
QY      289  DKLSAPYDQNSVYAGKAAKEALQAEGLPVDPTAPLFAFGRLEEQKGVDIILAALPK 348
Db      330  DKCIPCHYSVDDL-SGRAKCKGALQKELGPIRPDVPILGFTIGRLDYQKGDILQIILP 388
QY      349  ILATPKVQIAITLGTGKAAEKLVAIGTKYGRKGVVFKFSAPLAHMLTAGADFMVPSR 408
Db      389  LM-REDVQFVMLGSGDPELEDWMRSTESIFKDKRGWGFSPVSHRITAGCDILMPSR 447
QY      409  FEPCGLIQLHAMHYGTPVPVASTGGGLVDTV 438
Db      448  FEPCGLNQLYAMQYGTVPVHATGGLRDTV 477

RESULT 11
US-08-941-445A-13
; Sequence 13, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanning
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 583 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-941-445A-13

```

```

Query Match      37.5%; Score 859; DB 3; Length 583;
Best Local Similarity 42.2%; Pred. No. 2.4e-81;
Matches 190; Conservative 77; Mismatches 145; Indels 38; Gaps 12;

QY      4  IVMVAAEVPWSKGGGLGVDVTGGLPIELVKRGRHVMYTIAPRY-----DQYADAWDTSVV 57
Db      95  IVFTVGEASPYAKSGGLDVCGLPVALAARGHRVVMVMPRYLNGTSDKNYANAFYTEKH 154

```


Sequence 8, Application US/08836567
Patent No. 6130367

GENERAL INFORMATION:

APPLICANT: Kossmann, Jens
APPLICANT: Springner, Franziska

APPLICANT: Abel, Gernot

TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES

TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC

TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: FISH & NEAVE

STREET: 1251 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10020

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/836,567

FILING DATE: 24-JUL-1997

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP95/04415

FILING DATE: 09-NOV-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE P 44 41 408.0

FILING DATE: 10-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: Haley Jr., James F.

REGISTRATION NUMBER: 27,794

REFERENCE/DOCKET NUMBER: AGREVO-4

TELEPHONE: 212-596-9000

TELEFAX: 212-596-9090

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 767 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-836-567-8

Query Match 36.1%; Score 826; DB 4; Length 767;

Best Local Similarity 42.3%; Pred. No. 1.le-77; Mismatches 146; Indels 48; Gaps 11;

Matches 191; Conservative 67;

QY 2 LDIVMAAEVAPWSKTGGLGDTGVTGGLPIELVKRGHRYMTIAPRYDQYADAWDTSV---V 57

DB 276 MNILIVASECAPWSKTGGLGDTGVTGGLPIELVKRGHRYMTIAPRYDQYADAWDTSV---V 335

QY 58 VDIMEKRVYTHSIKKGVHVRWIDHPWFLAKVNGKTGSKLYGPRSGADYLDNHRKRALFC 117

DB 336 VDGQDVEVTYFQAFIDGVDFEVID----SHMERHIGNNIY----GGRNVDILKRMVLF 386

QY 118 KALEARVLP----FGPGEDCVFVANDHSAVLPVLLKDEYQPKGFTKAKSVLAHN 172

DB 387 KALEARVLP----FGPGEDCVFVANDHSAVLPVLLKDEYQPKGFTKAKSVLAHN 172

QY 173 IAFQGRWEAFKDKLPFAAFDKLAFSDGYAKVYVTEATPMEDEKPLTKYKKNWL 232

DB 446 IAHQGRGLEDFSVVDLPPHYMDFKLYD-----PVGGEHF---NIF 484

QY 233 KGLIADKLVTVSPNTATEIADAAGVDELTVIRAKG--LEGIVNGMDIEFWNPXTDK 290

DB 485 AAGLKTADRVVTVSHGYSWEIKT--SQGGWGLHQIINENMKLQIVNGIDTKWNPEDLV 543

QY 291 FLAP-----YDONSIVAGKAAAKALQALGLPVDPTAPLFAFTGRLEEQKGVDIILAAL 346

DB 544 HLQSDGYMNYSLDTLQTKFKPCKAAALOKELGLPVRDDVPLIGFTIGRIDPOKGVDLIAEAS 603

QY 347 PKILATPKVQITAILGTGKAAEYKIVNAIGTKYGRAGKGVVVKFSAPLAHMLTAGADEMLVP 406

DB 604 AWMWG-QDVQLVGLTGTGRDLQMLRQFECQHNDKIRGWGFSVKTSRITAGADILLMP 662

QY 407 SREPCGLIQLHAMHYGTVPVAVSTGGGLVDTV 438

DB 663 SREPCGLNQLYAMKYGTIPVHAVGGIRDTV 694

RESULT 15

US-09-196-390-6

Sequence 6, Application US/09196390

Patent No. 6307125

GENERAL INFORMATION:

APPLICANT: Block, Martina

APPLICANT: Lorz, Horst

APPLICANT: Lutticke, Stephanie

APPLICANT: Walter, Lennart

APPLICANT: Froberg, Claus

APPLICANT: Kossmann, Jens

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES

TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH

TITLE OF INVENTION: SYNTHESIS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave

STREET: 1251 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10020

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/196,390

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 196 21 588.9

FILING DATE: 29-MAY-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 196 36 917.7

FILING DATE: 11-SEP-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP97/02793

FILING DATE: 28-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Haley, Jr., James F.

REGISTRATION NUMBER: 27,794

REFERENCE/DOCKET NUMBER: AGREVO-9

TELEPHONE: (212) 596-9000

TELEFAX: (212) 596-9090

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 799 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-196-390-6

Query Match 35.7%; Score 817; DB 4; Length 799;

Best Local Similarity 41.8%; Pred. No. 1.le-76;

Matches 190; Conservative 67; Mismatches 144; Indels 54; Gaps 12;

QY 2 LDIVMAAEVAPWSKTGGLGDTGVTGGLPIELVKRGHRYMTIAPRYDQYADAWDTSV--VVD 59

DB 308 MNVVAECSPWCKTGGTGLGDTGGLGDTGGLGDTGGLGDTGGLGDTGGLGDTGGLGDTGGL 367

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 14:56:21 ; Search time 13.5267 Seconds
(without alignments)
3342.965 Million cell updates/sec

Title: US-09-980-771a-7

Perfect score: 2288

Sequence: 1 ALDIWVAEAPWSKTGGL.....AMHYGTVDPVASTGLVDTV 438

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1267	55.4	636	9	US-10-138-075-4
2	1247.5	54.5	609	9	US-10-138-075-2
3	1211.5	53.0	599	9	US-10-138-075-5
4	862	37.7	671	10	US-09-952-677-2
5	817	35.7	799	10	US-09-952-677-6
6	528.5	23.1	909	9	US-10-163-214-2
7	513.5	22.4	915	9	US-10-163-214-6
8	511.5	22.4	874	9	US-10-163-214-13
9	511.5	22.4	914	9	US-10-163-214-12
10	387	16.9	474	12	US-10-007-693-99
11	383.5	16.8	476	12	US-10-007-693-69
12	348	15.2	293	9	US-10-163-214-10
13	215.5	9.4	117	10	US-09-739-438-2
14	176.5	7.7	409	9	US-09-738-626-4738
15	164	7.2	143	10	US-09-739-438-4
16	156	6.8	381	9	US-09-738-626-5896
17	126	5.5	191	10	US-09-924-358-22
18	122.5	5.4	418	9	US-09-738-626-3951
19	110.5	4.8	121	9	US-10-163-214-4

20	109.5	4.8	387	9	US-09-738-626-3890	Sequence 3890, Ap
21	105.5	4.6	490	9	US-10-218-743-41	Sequence 41, Appl
22	105.5	4.6	509	9	US-10-218-743-35	Sequence 35, Appl
23	105.5	4.6	509	9	US-10-218-743-38	Sequence 38, Appl
24	104.5	4.6	2993	9	US-09-738-626-6239	Sequence 6239, Ap
25	103.5	4.5	385	10	US-09-767-041-18	Sequence 18, Appl
26	98.5	4.3	379	9	US-09-738-626-6655	Sequence 6655, Ap
27	97	4.3	489	10	US-09-815-242-11953	Sequence 11953, A
28	96.5	4.2	636	9	US-09-738-626-5590	Sequence 5590, Ap
29	96.5	4.2	216	9	US-10-222-162-46	Sequence 46, Appl
30	96.5	4.2	216	9	US-10-143-024-46	Sequence 46, Appl
31	93.5	4.1	732	10	US-09-381-624A-3	Sequence 3, Appl
32	92	4.1	195	10	US-09-815-028-8	Sequence 8, Appl
33	92	4.0	567	9	US-10-061-269-16	Sequence 16, Appl
34	91.5	4.0	536	9	US-10-218-743-21	Sequence 21, Appl
35	91.5	4.0	555	9	US-10-218-743-15	Sequence 15, Appl
36	91.5	4.0	555	9	US-10-218-743-18	Sequence 18, Appl
37	91.5	4.0	590	10	US-09-764-864-981	Sequence 981, App
38	91	4.0	1164	10	US-09-870-122-1	Sequence 1, Appl
39	90.5	4.0	249	10	US-09-765-272-6	Sequence 6, Appl
40	90.5	4.0	250	10	US-09-765-272-226	Sequence 226, App
41	90.5	4.0	266	9	US-09-769-787-165	Sequence 165, App
42	90.5	4.0	372	10	US-09-815-242-5277	Sequence 5277, Ap
43	90.5	4.0	384	10	US-09-815-242-12605	Sequence 12605, A
44	90.5	4.0	748	9	US-09-272-975-2	Sequence 2, Appl
45	90.5	4.0	753	9	US-09-272-975-58	Sequence 58, Appl

ALIGNMENTS

RESULT 1

US-10-138-075-4
; Sequence 4, Application US/10138075
; Publication No. US20030087369A1
; GENERAL INFORMATION:
; APPLICANT: Broglie, Karen E.
; APPLICANT: Butler, Karlene H.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Orozco, Emil M.
; TITLE OF INVENTION: Granule-Bound Starch Synthase
; FILE REFERENCE: BB1474 NA
; CURRENT APPLICATION NUMBER: US/10138,075
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/288,315
; PRIOR FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 636
; TYPE: PRT
; ORGANISM: Glycine max
US-10-138-075-4

Query Match 55.4%; Score 1267; DB 9; Length 636;
Best Local Similarity 55.4%; Pred. No. 1e-102;
Matches 251; Conservative 55; Mismatches 113; Indels 34; Gaps 7;

QY	2	LDIVMVAEAPWSKTGGLDVTGGLPTELVEGRHVMIIAPRYDQYADAWTSSVVDI- 60
Db	106	MTFIIGTEVAPWCVTGGDLGVLGPPALAGFGRHVMITVPRYDQYADAWTSSVVIK 165
QY	61	---MGEKRYFHSIKKGVRVWIDHPWFLAKVWGKTGSKLYGPRSGADYLDNHRKFLFC 117
Db	166	VCDRTKRYFFHCYKRGVDRVFDHPWFLKWKVGTGKLYGTGNGYEDNQLRSLFC 225
QY	118	KAATEAARVLPF-----GP-GEDCVFVANDWHSALVPVLLKDEYQPKGQFTKAKSVIA 169
Db	226	QAALAPRVLSLNSSKSYFSGPYGEDVIFVANDWHTALIPCYLKSVMYQSRGIYTNARVVC 285
QY	170	IHTAFQGRMWEAFKDKLFOAARDKIAFSDGYAKYVTEATPMEDEKPPUTGTYKKI 229

		Query Match	Score 817;	DB 10;	Length 799;
		Best Local Similarity	41.8%;	Pred. No. 4.3e-63;	
		Matches 190;	Conservative 67;	Mismatches 144;	Indels 54; Gaps 12;
QY	231 WLKGGIIADKLVTVSPNYATEIADAAGGVELDTVI--RAKGTEGIVNGMDIEBNPKT 288	: :: ::: : :	:	:	:: :: ::
Dd	276 FLKGAVTVADRITVTSQGSYEHTP-AEGGUGLNELLSRRKSVLINGVINGDIDNDNPPT 334				


```
Query Match 16.9%; Score 387; DB 12; Length 474;
Best Local Similarity 28.6%; Pred. No. 1.1e-25;
Matches 134; Conservative 63; Mismatches 182; Indels 90; Gaps 18;

QY 2 LDIVMAAEVAPWSKGGGLGDDVGGPIELVVKGRHVMVTIAPRYDQVADAWDSVVDIM 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MKIHTAIEFAPVKGAGLGLDALYGL-AKALAAHHTTEVVIPYKPLTFLPKEODLCSI- 58
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 GEKRYE-----HSIKGVHRVWIDHPWFLAKVWGKSGKLYGPRSGADYL 107
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 -QKLSYFFAGEQEATAESFYEGIKVTLFKL-----DPOPELFENAEITYS 104
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 108 DNHRKRALFCKAAI-----BAARVLPFGGDCVVFVANDWHSALVPVLLKDEYQPKGQFT 162
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 105 DDAFRFCFAFAAAASYIQEGGANIVHL-----HDWHTGLVAGLLKQ-QPCSQLQ 152
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 163 KAKSVLAHNIAPQGRWEEAFKDTKLPOAFDKLAFSDGYAKYVTEATPMEDEKPLT 222
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 153 KI--VLTNHFVGYGTYTTRILEASSLNE-----FYISOYQLFRDPQTCVL 196
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 223 GKYKKNLWKGGLIAADKLVTSPPNYATEIAADAAGVVELDTVIRAK--GIEGIVNGMD 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 197 -----LKGALYCSDFVTTSPTYAKELLEDYS-DYELHDAITARQHHLRGILNGID 246
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 281 IEWNPKTDKFLSAPYDQ-----NSVYAGKAAAKAELQAEGLPVPDPTAPLFAFIRGLE 334
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 247 TTWGPETDPLAKNYTKELFETPSIFFEAKAENKALYERLGLSLE-HSPVCVCIISRIA 305
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 335 EOKGVDIILALPKILATPKVQJAILGT--GKAAYEKLNV--AIGTKYKGRAGVVKFS 389
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 306 EQGPHPMKQAILHALENAVYTLI-IIGTCYGNQHEEFANLOESLANSPDVRI--LITYS 362
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 390 APLAHLMTAGADEMLVPSRFPGLIQLHAMHYGTPVVAAGVSTGGGLVDTV 438
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 363 DVLARQIFAADMICIPSMEEPCGLTQMIGNYGTVPVLRATGGGLADTV 411
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
US-10-007-693-69
; Sequence 69, Application US/10007693
; Patent No. US20020146776A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515C2
; CURRENT APPLICATION NUMBER: US/10/007.693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 69
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-007-693-69

Query Match 16.8%; Score 383.5; DB 12; Length 476;
Best Local Similarity 29.3%; Pred. No. 2.2e-25;
Matches 138; Conservative 62; Mismatches 172; Indels 99; Gaps 21;

QY 4 IWMVAEAPWSKGGGLGDDVGGPIELVVKGRHVMVTIAPRY--DQVADAWDSVVDI 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 IVQVAVEFTPIVKVGGGLGDAVASLAKELAKQ-NDVEVLLPHYLPSKFS--QV 54
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 MGEKRYEHSIKGVHRVWIDHPWFLAKVWGKSGKLYGP-----RSGAD 105
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 55 LSERSFYE-----FLGKQASALSYSEGLTLITITLDSQIELEFSTISV 99
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 106 YLDNH-KRFALFCKAAIEARVLPFGGDCVVFVANDWHSALVPVLLKDEYQPKGQFTKA 164
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 100 YSENNVVRFSFAFAAA--AAVQLQADPAD--IVHLHDWVHVLGAGLLKNPLNP----VHS 151
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 165 KSVLAHNIAPQGRWEEAFKDTKLPOAFDKLAFSDGYAKYVTEATPMEDEKPLTCK 224
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Query Match 15.2%; Score 348; DB 9; Length 293;
Best Local Similarity 37.1%; Pred. No. 1.4e-22;
Matches 82; Conservative 44; Mismatches 77; Indels 18; Gaps 5;

QY 228 KINWLKGGIIAADKLVTSPPNYATEIAADAAGVVELDTVIRAKGIEGIVNGMDIEWNPK 287
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 RINAVGAVVYSNIVTTSPTYALEVRSEGGGLQDTLVHRSKRLGILNGIDTDTWNP 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 288 TDKELSAFYDQNSVYAGKAAAKAELQAEGL-PVDPPTAPLFAFIRGLEQKGVDIILAL 346
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 TDRLKLVQYNKDL-QGKAANKAALREQLNLASAYPSQPLVGCITRLVAQKGVHLIRHAI 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 347 PKILATPKVQJAILGTGKAAYEKLNVATGTYKGR-----KGVVKFSAPLAHMT 397
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 YK-TAELGGQFVLLGSSP-----VPEIQREFEGADHFNQNNNRLILIKYDDALSHCIY 172
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 398 AGADFMVPSRFPGLIQLHAMHYGTPVVAAGVSTGGGLVDTV 438
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 173 AASDMFTVPSIFEPGCLTQMIAIRYGVPIVEKGTGLNDSV 213
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
US-09-739-438-2
; Sequence 2, Application US/09739438
; Patent No. US2002029394A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Beckles, Diane M.
; APPLICANT: Thorpe, Cathy
; TITLE OF INVENTION: Homologs of Starch Synthase D01
```

```

RESULT 14
US-09-738-626-4738
; Sequence 4738, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIYOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHITO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDE
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/37748
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/15916
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/28098
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4738
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4738

```

Query Match	7.7%;	Score 176.5;	DB 9;	Length 409;
Best Local Similarity	27.6%;	Pred. No. 2,6e-07;		
Matches	66;	Conservative 37;	Mismatches 99;	Indels 37; Gaps 11;
210	ATPWEDE--KPLTGTCTYKYNWL-KGIIAADKLVTSPNYATEIAADAAGGVELDTV	266		
129	AHSLDPDPWKREQLGGYDVSSSEKNAMEYADAVINAVSRMKDLSLA-AYPIEDPNV	187		
267	IRAKGLEGIINGMDIEENPKTDKFLSAPYQDONSVIAGKAAKALQELGPLVDPTAPL	326		
188	-----RVVLNGIDTLMQPR-----PTFDD-----AEDSVLSLG--VDPORPI	224		
327	FAFIGLEEQKGDVILIILALPKILATPKVQITAILGTGKAAYE--KLNVNATGTYTKGRAGK	384		

```

Query Match      7.2%; Score 164; DB 10; Length 143;
Best Local Similarity 23.9%; Pred. No. 7.1e-07;
Matches 49; Conservative 23; Mismatches 71; Indels 62; Gaps 5;

106 YLDMHKREAFCKAAIEAARVLFFPGEDCVFVANDWHSALYVLLKDEYQPKGFYKAK 165
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 YSNDESFFGFCHSALEFLRQNGSSPD---IITHDWSSAPVAVLFEKQYQANG-LSNGR 56
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

166 SVLAIHNAFQGRWEEAFKDTKLQPAQFDFKLAFSQYAKVYTEATPWEDEKPLTQKT 225
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
57 VVFTIHNLEFGAHH-----IGKA 74

226 YKKINLWKGGLIIRADKLVTYSPNYATEIARAAGGVVELDTVIRAKIGIYNGMDIERNN 285
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
75 MAR-----CDKATVTYTYTSREVSHGAIAPHSX-----FHGIRNGIDPDWD 118

286 PKTDKFLSAPYDQNSVYAGKAAAKE 310

119 PYSDNFIPIVHYTSENVTYXGKSAKK 143

```

Search completed: June 4, 2003, 15:15:32
Job time : 20.5267 secs

This Page Blank (uspto)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2003, 08:28:25 ; Search time 2768.82 Seconds
(without alignments)
16743.883 Million cell updates/sec

Title: US-09-980-771A-8

Perfect score: 1593

Sequence: 1 gcgtggacatcgtgatgtt.....ccgttgccgagaagatcccc 1593

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1593	100.0	1593	6	AX049329	Sequence
2	1593	100.0	1593	6	AX049325	Sequence
3	1593	100.0	2124	6	AX049323	Sequence
4	1593	100.0	3117	6	AX049322	Sequence
5	1591.4	99.9	3240	8	AF026420	Chlamydom
6	1314	82.5	1314	6	AX049327	Sequence
7	603.8	37.9	5856	8	AF433156	Chlamydom
8	535.8	33.6	2289	8	AF286320	Triticum v
9	530	33.3	1812	8	AF486514	Hordeum v
10	530	33.3	2311	8	HVWAXYR	Barley mRNA
11	528.4	33.2	1827	8	AF486515	Hordeum v
12	528.4	33.2	1827	8	AF486518	Hordeum v
13	525.2	33.0	1812	8	AF486519	Hordeum v
14	525.2	33.0	1827	8	AF486516	Hordeum v
15	525.2	33.0	1827	8	AF486517	Hordeum v
16	524.6	32.9	2127	8	AF163319	Triticum
17	523	32.8	1605	8	AF250137	Triticum
18	505.4	31.7	2028	8	AF113844	Triticum
19	505.2	31.7	2186	8	TAWAXYS	Wheat waxy
20	505.2	31.7	2267	8	AY050174	Triticum
21	481.4	30.2	1801	8	AF113843	Triticum
22	474.8	29.8	2267	6	AR195560	Sequence
23	472.8	29.7	2115	8	AB089141	Setaria i
24	453.2	28.4	2287	8	SB023945	Sorghum bic
25	449	28.2	1937	8	AB066093	Oryza sat
26	449	28.2	1937	8	AB066094	Oryza sat
27	447.4	28.1	2542	6	AR106491	Sequence
28	447.4	28.1	2542	8	OSWX	O.sativa wa
29	347	21.8	2168	8	MEGBSS	M.esculenta
30	340.2	21.4	2270	8	AF210699	Perilla f
31	329.6	20.7	2221	8	AB029546	Phaseolus
32	321.6	20.2	2211	8	AB071604	Ipomoea b
33	316.8	19.9	2161	6	AX394246	Sequence
34	313.2	19.7	2345	8	AMA6293	Antirrhin
35	312.8	19.6	3049	8	AF026421	Chlamydom
36	305.6	19.2	2124	8	IBU41126	Ipomoea bat
37	298	18.7	2092	8	AY094405	Arabidops
38	298	18.7	2115	8	AY088544	Arabidops
39	298	18.7	2192	8	AY123983	Arabidops
40	289.6	18.2	2107	8	AF097922	Astragalu
41	283.2	17.8	764	8	AF446083	Miscanthu
42	279.6	17.6	2035	8	PSSTARSYN	P.sativum m
43	274.2	17.2	2081	8	AF109395	Triticum
44	274	17.2	764	8	AF446082	Microsteg
45	273.2	17.2	721	8	AF446081	Cleistach

ALIGNMENTS

RESULT 1	AX049329	Sequence 8	from Patent WO0071734.	DNA	Linear	PAT 12-JAN-2001
LOCUS	AX049329	Sequence 8	from Patent WO0071734.	DNA	Linear	PAT 12-JAN-2001
DEFINITION	AX049329	Sequence 8	from Patent WO0071734.	DNA	Linear	PAT 12-JAN-2001
ACCESSION	AX049329	Sequence 8	from Patent WO0071734.	DNA	Linear	PAT 12-JAN-2001
VERSION	AX049329.1	GI:12226096				
KEYWORDS		synthetic construct.				
SOURCE		synthetic construct				
ORGANISM		artificial sequences.				
REFERENCE		1 (bases 1 to 1593)				
AUTHORS		D'Hulst, C. and Ball, S.				
TITLE		Starch granules containing a recombinant polypeptide of interest,				
JOURNAL		method for obtaining same and uses				
		Patent: WO 0071734-A 8 30-NOV-2000;				

Pred. No. is the number of results predicted by chance to have a

```

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)
FEATURES
  Location/Qualifiers
    1..1593
      /organism="synthetic construct"
      /db_xref="taxon:32630"
      /note="fragment de l'ADN complet codant la GBSSI de
      Chlamydomonas reinhardtii"
      <1..>1593
    CDS
      /note="unnamed protein product"
      /codon_start=1
      /trans_table=11
      /protein_id="CAC21783.1"
      /db_xref="GI:12226097"
      /translation="ALDIVMVAEVAEAPWSKTTGGLDVTGGLPIELVRRHVRMTIAPR
      YDOADAWDSVVVDIMGEVRYPHSKVGRVWIDHPFLAKWVKTKGSKLYGPRS
      GADYLDNKEKFAKKAIEAARVLPFGEDQCVFANDWHSALVPLLKDEYOPKGO
      FTKAKSVLAHNTAFQGRMEEAFKDKLQAADFKAFLSDGYAKYVTEAPMEDEK
      PPLGKTKIKLWKGGLIAADKLVTSPNTATEIAADAGVELDVIIRAKLEGIV
      NGMDIEWNPRTDFELSAFYDONSVYAKAAKEALQAEGLPDPVAPLFAFIEGRDE
      EQGVDIILALPKILATPKVQIAILGTAKAAYEKLNIAIGTKYKRAKGVYFTSAPL
      AHMLTAGADFMVPSRFPFCGLIOLHAMHYGVTPVASTGGGLVTVKRGVTFHMGAL
      NPDKLDEADADALAATVRRASEVEFAGGRYPBMWANCISQDLWSKPAQKWEGLLEEVV
      YKGKGVATAKKEELKVPVAEKIP"
      290 a 539 c 510 g 254 t
BASE COUNT
ORIGIN
  Query Match 100.0%; Score 1593; DB 6; Length 1593;
  Best Local Similarity 100.0%; Pred. No. 2.4e-166;
  Matches 1593; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGCTGGACATCGTATGGTTGCTGCTGAGGTGCGCCCTTGTTCCAAAGACGGCGGCCGTG 60
DB 1 GCGCTGGACATCGTATGGTTGCTGCTGAGGTGCGCCCTTGTTCCAAAGACGGCGGCCGTG 60
QY 61 GCGATGTGACTGGTGGCTGCTATATGAGTGGTCAAGCGGGCCACCGCGTCATGACC 120
DB 61 GCGATGTGACTGGTGGCTGCTATATGAGTGGTCAAGCGGGCCACCGCGTCATGACC 120
QY 121 ATTGCCCTCGCTAGCAGCAGTACGCTGACGCTGGGACACCTCGGTGGTGGTGGACATC 180
DB 121 ATTGCCCTCGCTAGCAGCAGTACGCTGACGCTGGGACACCTCGGTGGTGGTGGACATC 180
QY 181 ATGGGGGAGAGGTGCGGTACTTCCACTCCATCAAGAGGGCGTGACCGCGTGTGGATT 240
DB 181 ATGGGGGAGAGGTGCGGTACTTCCACTCCATCAAGAGGGCGTGACCGCGTGTGGATT 240
QY 241 GACCAACCTGCTGCTTGCCCAAGTCTGCGGCAAGACCGGCTCCAAAGCTGTACGGCCCC 300
DB 241 GACCAACCTGCTGCTTGCCCAAGTCTGCGGCAAGACCGGCTCCAAAGCTGTACGGCCCC 300
QY 301 CGCTCCGGCGCTGACTACTGGACAAACCAAGCGCTTCGCCCTGTCTGCAAGGCCGT 360
DB 301 CGCTCCGGCGCTGACTACTGGACAAACCAAGCGCTTCGCCCTGTCTGCAAGGCCGT 360
QY 361 ATTGAGCTGCGCGGTGCTGCTGCTGCGCCCGGGAGAGTGCCTCTCTGTTGGCCAC 420
DB 361 ATTGAGCTGCGCGGTGCTGCTGCTGCGCCCGGGAGAGTGCCTCTCTGTTGGCCAC 420
QY 421 GACTGGCACTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 GACTGGCACTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 TTCACAAAGGCGAAGTGGTGGTATCCACAAATCCGCTTCACAAAGTGGCTTCTCG 540
DB 481 TTCACAAAGGCGAAGTGGTGGTATCCACAAATCCGCTTCACAAAGTGGCTTCTCG 540
QY 541 GAGGAGGCTTTCAGACACAGAGTGCCTGCGCGCGCGCTTTCACAAAGTGGCTTCTCG 600
DB 541 GAGGAGGCTTTCAGACACAGAGTGCCTGCGCGCGCGCTTTCACAAAGTGGCTTCTCG 600
QY 601 GAGCGGTATGCCAAGTGTACACTGAGGCGCACCCCATGAGGAGGACGAGAGCCCGCCG 660
DB 601 GAGCGGTATGCCAAGTGTACACTGAGGCGCACCCCATGAGGAGGACGAGAGCCCGCCG 660
```

```

QY 661 CTGACGGGAAGACCTACAAGAAGATCAACTGGCTGAAGGGTGGCATTTATCCGCCCGGAC 720
DB 661 CTGACGGGAAGACCTACAAGAAGATCAACTGGCTGAAGGGTGGCATTTATCCGCCCGGAC 720
QY 721 AAGCTGGTACTGTGTGCGCCAACTACGCGACCGAGATCGTCCGATGCCCGCGCGGT 780
DB 721 AAGCTGGTACTGTGTGCGCCAACTACGCGACCGAGATCGTCCGATGCCCGCGCGGT 780
QY 781 GTGAGCTGGACACCGTCAATCCGCGCAAGGCATTGAGGGCATTTGTAACGGCATGGAC 840
DB 781 GTGAGCTGGACACCGTCAATCCGCGCAAGGCATTGAGGGCATTTGTAACGGCATGGAC 840
QY 841 ATTGAGGAGTGAACCCCAAGCCCGCCCAAGAGTTCCTGTCTGCGCCCTACGACCAACACG 900
DB 841 ATTGAGGAGTGAACCCCAAGCCCGCCCAAGAGTTCCTGTCTGCGCCCTACGACCAACACG 900
QY 901 GTCTAGCGGGCAAGGCCCGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
DB 901 GTCTAGCGGGCAAGGCCCGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
QY 961 GACCCACACCGCCCGCTGTTGCTTCATCGCGCGCTGAGGAGGAGGAGGAGGAGGAGGAG 1020
DB 961 GACCCACACCGCCCGCTGTTGCTTCATCGCGCGCTGAGGAGGAGGAGGAGGAGGAGGAG 1020
QY 1021 ATCATCTGGCGCCCTGCGCAAGATCCTTGCGCACCCCAAGGTGAGATCGCCATCCTCTG 1080
DB 1021 ATCATCTGGCGCCCTGCGCAAGATCCTTGCGCACCCCAAGGTGAGATCGCCATCCTCTG 1080
QY 1081 GGTACCGCAAGCGCCCTACGAGAAGCTGTGTAAGCCATCGGACCAAGTACCAAGGC 1140
DB 1081 GGTACCGCAAGCGCCCTACGAGAAGCTGTGTAAGCCATCGGACCAAGTACCAAGGC 1140
QY 1141 CGCGCAAGGGGTGTCAAGTTCCTGCGCGCCCTGCGGCACATGCTACCGCGCGCC 1200
DB 1141 CGCGCAAGGGGTGTCAAGTTCCTGCGCGCCCTGCGGCACATGCTACCGCGCGCC 1200
QY 1201 GACTTCATGCTGCGCTGCGCGCTGCGAGCCTGCGCGCTGCGAGCTGCGAGCTGCGAGCATG 1260
DB 1201 GACTTCATGCTGCGCTGCGCGCTGCGAGCCTGCGCGCTGCGAGCTGCGAGCTGCGAGCATG 1260
QY 1261 CACTAGGTACCGTGGCGGTGAGCTCCACCGCGCGCTGCGAGCCTGCGAGCTGCGAGCATG 1320
DB 1261 CACTAGGTACCGTGGCGGTGAGCTCCACCGCGCGCTGCGAGCCTGCGAGCTGCGAGCATG 1320
QY 1321 GCGCTCACCGGTTCACATGCGCGCCCTGAACCCCGCAAGCTGAGCAGGAGTGAAGGC 1380
DB 1321 GCGCTCACCGGTTCACATGCGCGCCCTGAACCCCGCAAGCTGAGCAGGAGTGAAGGC 1380
QY 1381 GAGCCCTGCGCGCCCGCTGCGCGCTGCGAGGAGTGTGCGGGCGCGCGCTACCCC 1440
DB 1381 GAGCCCTGCGCGCCCGCTGCGCGCTGCGAGGAGTGTGCGGGCGCGCGCTACCCC 1440
QY 1441 GAGATGTTGSCCAACTGCTACCGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
DB 1441 GAGATGTTGSCCAACTGCTACCGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
QY 1501 GAGGGCTGCTGAGGAGGAGTGTGTAAGGCAAGGGGCGGCTGCGCCACGAGGAGGAG 1560
DB 1501 GAGGGCTGCTGAGGAGGAGTGTGTAAGGCAAGGGGCGGCTGCGCCACGAGGAGGAG 1560
QY 1561 GAGATCAAGGTGCGCGCTGCGGAGAGATCCCC 1593
DB 1561 GAGATCAAGGTGCGCGCTGCGGAGAGATCCCC 1593
RESULT 2
AX049325 Sequence 4 from Patent WO0071734. linear PAT 12-JAN-2001
LOCUS
DEFINITION
ACCESSION AX049325
VERSION AX049325.1 GI:12226092
KEYWORDS
```



```
|||||
Db 1561 GAGATCAAGGTGCCGTTGCCGAGAGATCCCC 1593
|||||
RESULT 3
AX049323
LOCUS AX049323
DEFINITION Sequence 2 from Patent WO001734.
ACCESSION AX049323
VERSION AX049323.1 GI:12226090
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
          artificial sequences.
REFERENCE 1 (bases 1 to 2124)
AUTHORS D'Hulst,C. and Ball,S.
TITLE Starch granules containing a recombinant polypeptide of interest,
        method for obtaining same and uses
JOURNAL Patent: WO 001734-A 2 30-NOV-2000;
FEATURES Location/Qualifiers
         1..2124
            /organism="synthetic construct"
            /db_xref="taxon:32630"
            /note="fragment de la sequence compl te de l'DNAc codant
            la GBSSI de Chlamydomonas reinhardtii"
         1..>2124
            /note="unnamed protein product"
            /codon_start=1
            /transl_table=11
            /protein_id="CAC21780.1"
            /db_xref="GI:12226091"
            /translation="MAVASTSRPSSARPIVINAASFVGVKKTANOLLRELARGSARKST
            SRGAVTCATCATCALDVMVAEAVPMSKGTGLDVTGGLPIELVKKGRHVMITAPRY
            DOYADADKTSVVDIMGEKRYRPHSIKKGVHRVWIDHFWFLAKVWGTSKLYGPRSG
            ADYLNDHRALNFCRKALEAARVLPFGPGEDCVFVANDWSALVPVLLKDEYQPKQF
            TKASVLALNIAKPGRMWEEAFKILPQAAEDFKLAFSDGYAKVYTEATPMEDEKRP
            PLTGKTKYKTNLWKGHTIAADKLIVTSPTYATEIADAAGGVELDTVIRAKGIEGINV
            GMDEENPNPTKFLSAPYQNSVYAGKAKEALQAEGLPVDPTAPLAFAGRLLEE
            OKGVDIILALPKLAPKVIQIAILGTGKAAYEKLVNAIGTKYGRKAGVVKFSAPLA
            HMLTAGADEMLVSRFPFPGCLGLIHAMHYGTVPVAVSTGGLVDIVKGVTFHMGALN
            PKLDEADALALATVRSAREVAGGRYPENVANCISQDLSWSKPAKQWGLLEEVY
            KGGVATATKEELKVPVAAEKIPGLDIPAVSTAPNLKLPVSAVSEVNGAARAPKVGTIATA
            MGAWRATTPGSPGPAATPKVITPKPALPATAPKPTAGLKLAGEASTTSUSENGAASN
            GNGNGASAKTSAAKPLVSAATRKSA"
BASE COUNT 369 a 773 c 656 g 326 t
ORIGIN
Query Match 100.0%; Score 1593; DB 6; Length 2124;
Best Local Similarity 100.0%; Pred. No. 2.2e-166;
Matches 1593; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGCTGGACATCGTGATGGTGGCTGCTGAGTTCGCCCTTGGTCCCAAGACGGCGGCGCTG 60
Db 172 GCGCTGGACATCGTGATGGTGGCTGCTGAGTTCGCCCTTGGTCCCAAGACGGCGGCGCTG 231
Qy 61 GCGGATGTGACTGGTGGCTGCTTATGAGTGTGTCATAGCGCGGCCACCGGCTCATGACC 120
Db 232 GCGGATGTGACTGGTGGCTGCTTATGAGTGTGTCATAGCGCGGCCACCGGCTCATGACC 291
Qy 121 ATTGCCCTCGCTAGGACCATGAGTACGCTACGCGCTGGGACACCTCGGTGGTGGACATC 180
Db 292 ATTGCCCTCGCTAGGACCATGAGTACGCTACGCGCTGGGACACCTCGGTGGTGGACATC 351
Qy 181 ATGGCGAGAGAGTCCGCTACTTCCATCCATCAAGAGGGCGGTGACCGGCTGTGGATT 240
Db 352 ATGGCGAGAGAGTCCGCTACTTCCATCCATCAAGAGGGCGGTGACCGGCTGTGGATT 411
Qy 241 GACCACCCCTGGTTCCTGGCCAAAGTCTGGGSCAAGACCGGCTCCAAAGCTGTACGGCCCC 300
Db 412 GACCACCCCTGGTTCCTGGCCAAAGTCTGGGSCAAGACCGGCTCCAAAGCTGTACGGCCCC 471
Qy 301 CGCTCGGGCGGTGACTACTTGGACAAACCAAGCGCTTCGCCCTGTCTGCAAGSCCGCT 360
|||||
472 CGCTCCGCGCTGACTACTTGGACAAACCAAGCGCTTCGCCCTGTCTGCAAGSCCGCT 531
361 ATTGAGGCTGCCGCGTCTGCCCTTCGGCCCGCGGAGGACTGGTCTTCGTGGCCCAAC 420
532 ATTGAGGCTGCCGCGTCTGCCCTTCGGCCCGCGGAGGACTGGTCTTCGTGGCCCAAC 591
421 GACTGGCAGCTCCGCGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
592 GACTGGCAGCTCCGCGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 651
481 TTCACCAAGGCGAAGTTCGGTCTGGCTATCCACAAATATCGCTTCCAGGGCGCATGTGG 540
652 TTCACCAAGGCGAAGTTCGGTCTGGCTATCCACAAATATCGCTTCCAGGGCGCATGTGG 711
541 GAGGAGGCTTTCAAGGACACAGGCTGCCCGCAGGCGCTTTTGACAAAGTGGCCTTCTCG 600
712 GAGGAGGCTTTCAAGGACACAGGCTGCCCGCAGGCGCTTTTGACAAAGTGGCCTTCTCG 771
601 GACGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
772 GACGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 831
661 CTGACGGGAAAGACCTACAAGAAGATCAACTGGCTGAAGGTGGCATTCATCGCCGCCGAC 720
832 CTGACGGGAAAGACCTACAAGAAGATCAACTGGCTGAAGGTGGCATTCATCGCCGCCGAC 891
721 AAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
892 AAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 951
781 GTGGAGTGGACACCGTTCATCGCCGCCAAGGCGCATTCAGGGCATTCAGGGCATTCAGG 840
952 GTGGAGTGGACACCGTTCATCGCCGCCAAGGCGCATTCAGGGCATTCAGGGCATTCAGG 1011
841 ATTGAGGAGTGAACCCCAAGACCGCACAAAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
1012 ATTGAGGAGTGAACCCCAAGACCGCACAAAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1071
901 GTCTACCGCGCAAGCGCGCCGCCAAGGAGGCGCTTCGAGGCGGAGCTGGGCGCTGGCTGTG 960
1072 GTCTACCGCGCAAGCGCGCCGCCAAGGAGGCGCTTCGAGGCGGAGCTGGGCGCTGGCTGTG 1131
961 GACCCACCGCGCCCGCTGTTTCGCTTCATCGCGCGCGCTGAGGAGGAGGAGGAGGAGGAG 1020
1132 GACCCACCGCGCCCGCTGTTTCGCTTCATCGCGCGCGCTGAGGAGGAGGAGGAGGAGGAG 1191
1021 ATCATCTTGGCGCGCTGCCCAGAGATTCCTGGCCACCCCAAGGTGCAGATCGCCATCTCTG 1080
1192 ATCATCTTGGCGCGCTGCCCAGAGATTCCTGGCCACCCCAAGGTGCAGATCGCCATCTCTG 1251
1081 GGTACCGCGCAAGCGCGCGCTACGAGAAGTGTGTAAGCGCATCGGCACCAAGTACAGGCG 1140
1252 GGTACCGCGCAAGCGCGCGCTACGAGAAGTGTGTAAGCGCATCGGCACCAAGTACAGGCG 1311
1141 GCGGCCAAGGCGGTGGTCAAGTTCCTGGCGCGCGCTGGCGCACATCTCACCGCGCGCG 1200
1312 CCGCGCAAGGCGGTGGTCAAGTTCCTGGCGCGCGCTGGCGCACATCTCACCGCGCGCG 1371
1201 GACTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
1372 GACTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1431
1261 CACTACGCTACCGTCCCGCTGGTAGCCTCCACCGGCGGCTGGTAGCCTCAGGAGGAG 1320
1432 CACTACGCTACCGTCCCGCTGGTAGCCTCCACCGGCGGCTGGTAGCCTCAGGAGGAG 1491
1321 GCGCTACCGGCTTCCACATGGGCGGCTGAAACCCGCAAGCTGACGAGCTGACGCG 1380
1492 GCGCTACCGGCTTCCACATGGGCGGCTGAAACCCGCAAGCTGACGAGCTGACGCG 1551
1381 GACGCGCTGGCGCGCACCGTGGCGCGCTGCGCAGCAGCTGTTTGGGCGCGCGCTACCG 1440
|||||
```


Db 1089 GTGGAGCTGACACCGCTATCCGCGCCAAAGGCGATGAGGGCATCTGTAACGCGATGGAC 1148
QY 841 ATTGAGAGTGAACCCCAAGACCGACAAAGTTCCTGCTCGGCCCTACACACGACAGC 900
Db 1149 ATTGAGAGTGAACCCCAAGACCGACAAAGTTCCTGCTCGGCCCTACACACGACAGC 1208
QY 901 GTCTAGCGCGGCAAGCGCCGCCCAAGAGGCGCTCGAGCGCGAGCTGGCGCTGCGCTGTG 960
Db 1209 GTCTAGCGCGGCAAGCGCCGCCCAAGAGGCGCTCGAGCGCGAGCTGGCGCTGCGCTGTG 1268
QY 961 GACCCCAAGCGCCGCCCTGTGCTGCTATCGCGCGCGCTGGAGGAGCAGAAGGTGTGGAC 1020
Db 1269 GACCCCAAGCGCCGCCCTGTGCTGCTATCGCGCGCGCTGGAGGAGCAGAAGGTGTGGAC 1328
QY 1021 ATCATCTCTGGCGCGCTCGCCCAAGATCTGGCCACCGCCCAAGGTGCAGATCCGCATCTTG 1080
Db 1329 ATCATCTCTGGCGCGCTCGCCCAAGATCTGGCCACCGCCCAAGGTGCAGATCCGCATCTTG 1388
QY 1081 GGTACCGGCAAGCGCGCTACGAGAAGCTGTGAACGGCATCGGCACCAAGTACAAGGCG 1140
Db 1389 GGTACCGGCAAGCGCGCTACGAGAAGCTGTGAACGGCATCGGCACCAAGTACAAGGCG 1448
QY 1141 CGCGCCAAAGGCGGTGCTCAAGTTCGCGCGCCCTGGCGCACATGCTACCGCGCGCGCC 1200
Db 1449 CGCGCCAAAGGCGGTGCTCAAGTTCGCGCGCCCTGGCGCACATGCTACCGCGCGCGCC 1508
QY 1201 GACTTCATCTGCTGTCGCGCTCGCGCTTCGAGCGCTCGCGCGCTGATCCAGCTGCACGCCATG 1260
Db 1509 GACTTCATCTGCTGTCGCGCTTCGAGCGCTCGCGCGCTGATCCAGCTGCACGCCATG 1568
QY 1261 CACTACGCTACGCTGCGCGCTGAGCTCCACCGCGCGCTGCTGCGACCGCTCAAGGAG 1320
Db 1569 CACTACGCTACGCTGCGCGCTGAGCTCCACCGCGCGCTGCTGCGACCGCTCAAGGAG 1628
QY 1321 GCGCTCACCGCTTCCATCATGGCGCGCTGAGCTCCACCGCGCGCTGCTGCGACCGCTCAAGGAG 1380
Db 1629 GCGCTCACCGCTTCCATCATGGCGCGCTGAGCTCCACCGCGCGCTGCTGCGACCGCTCAAGGAG 1688
QY 1381 GACGCGCTGCGCGCGCTGCGCGCTGCGCGCGCTGCTGCGAGCGAGTGTTCGCGCGCGCTACCGCC 1440
Db 1689 GACGCGCTGCGCGCGCTGCGCGCTGCGCGCGCTGCTGCGAGCGAGTGTTCGCGCGCGCTACCGCC 1748
QY 1441 GAGATGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCTGCGAGCGAGTGTTCGCGCGCGCTACCGCC 1500
Db 1749 GAGATGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCTGCGAGCGAGTGTTCGCGCGCGCTACCGCC 1808
QY 1501 GAGGCGCTGCTGAGGAGGTGTGACGCAAGGCGCGCTGCGCGCGCTGCGCGCGCTACCGCGAGGAG 1560
Db 1809 GAGGCGCTGCTGAGGAGGTGTGACGCAAGGCGCGCTGCGCGCGCTGCGCGCGCTACCGCGAGGAG 1868
QY 1561 GAGATCAAGTGTGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCT 1593
Db 1869 GAGATCAAGTGTGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCT 1901

RESULT 6
AX049327
LOCUS AX049327 1314 bp DNA linear PAT 12-JAN-2001
DEFINITION Sequence 6 from Patent WO0071734.
ACCESSION AX049327
VERSION AX049327.1 GI:12226094
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 1314)
AUTHORS D'Huist,C. and Ball,S.
TITLE Starch granules containing a recombinant polypeptide of interest,
method for obtaining same and uses
JOURNAL Patent: WO 0071734-A 6 30-NOV-2000;
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)
FEATURES
Location/Qualifiers
1..1314
source

/organism="synthetic construct"
/db_xref="taxon:32630"
/note="fragment de l' DNAC complet codant la GBSSI de
Chlamydomonas reinhardtii"
<1..>1314
/codon_start=1
/transl_table=11
/protein_id="CAC21782.1"
/db_xref="GI:12226095"
/translation="ALDIDVMAAEVAPWSKTGGLGDTGGTGLPIELVKRGRHYMTIAPR
YDOYADWDTSVVDLMGEKRYFHSIKKGVHRWIDHPWLAKVWGKTSKLYGPRS
GADYLDNHRFALFACALEAREAVLPFGGEDCVFANDWHSALYVLLKDYQKQK
FTKASVLAIHIAFQGRWEAFKDTLPQAFDLAFSDGAYKYTEATPMEDEK
PPLTKYTKKINLWGLADLVTSVNTATEIAADAAGVELDTIRAKIGIV
NGMDLEWNPKTDFLSAPDONSVAKAAKEALQALGLPVPYPTAPLAFIGRIE
FOKVDYIIAALPKLAPKVQIALLTGKAYEKLVAIGTGYKGRAGVVKFSAPL
AHMLTAGADEMLVPSRFPFCGLIQHLHAMHYGVFVVASITGGLVDIV"
BASE COUNT 238 a 449 c 408 g 219 t

Query Match 82.5%; Score 1314; DB 6; Length 1314;
Best Local Similarity 100.0%; Pred. No. 1.le-135;
Matches 1314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGCTGGACATCGTGATGTTGCTGCTGAGTGCGCCCTTGTGCTCCAGAGCGCGCGCTG 60
Db 1 GCGCTGGACATCGTGATGTTGCTGCTGAGTGCGCCCTTGTGCTCCAGAGCGCGCGCTG 60
QY 61 GCGGATGTGACTGGTGGCTGCTATTGAGCTGGTCAAGCGCGCCACCGCGCTCATGACC 120
Db 61 GCGGATGTGACTGGTGGCTGCTATTGAGCTGGTCAAGCGCGCCACCGCGCTCATGACC 120
QY 121 ATTGCGCGCTGCTACGACCAAGTACGCTGAGCGCTGGGACACCTCGGTGCTGCTGGACATC 180
Db 121 ATTGCGCGCTGCTACGACCAAGTACGCTGAGCGCTGGGACACCTCGGTGCTGCTGGACATC 180
QY 181 ATGGGCGAGAGGTCGGCTACTTCCACTCCATCAAGAGGGGTGACCGCGCTGGATT 240
Db 181 ATGGGCGAGAGGTCGGCTACTTCCACTCCATCAAGAGGGGTGACCGCGCTGGATT 240
QY 241 GACCAACCGCTGCTGCGCAAGGTCTGGGCAAGCGCTGCGCGCTGCGCGCTGACGCGCC 300
Db 241 GACCAACCGCTGCTGCGCAAGGTCTGGGCAAGCGCTGCGCGCTGCGCGCTGACGCGCC 300
QY 301 CGCTCCGCGCTGCTACTACCTGGACAAACCAAGCGCTTCCGCGCTGCTTCTGCAAGCGCGT 360
Db 301 CGCTCCGCGCTGCTACTACCTGGACAAACCAAGCGCTTCCGCGCTGCTTCTGCAAGCGCGT 360
QY 361 ATTGAGGCTGCGCGCGCTGCTGCGCGCTGCGCGCGCGGAGGACTGCTGCTGCGCGAAC 420
Db 361 ATTGAGGCTGCGCGCGCTGCTGCGCGCTGCGCGCGCGGAGGACTGCTGCTGCGCGAAC 420
QY 421 GACTGGCACTCGCGCGCTGCTGCGCGCTGCTGCTGAGGAGGAGTACCGCGCGCGCGAG 480
Db 421 GACTGGCACTCGCGCGCTGCTGCGCGCTGCTGCTGAGGAGGAGTACCGCGCGCGCGAG 480
QY 481 TTCACCAAGGCGCAAGTTCGGTGTGCTGCTATCCACAAACATCGCTTCCAGGCGCGCTG 540
Db 481 TTCACCAAGGCGCAAGTTCGGTGTGCTGCTATCCACAAACATCGCTTCCAGGCGCGCTG 540
QY 541 GAGGAGGCTTTCACGAGCACCAAGTGCCTGCGCGCGCGCGCTTGTGACAACTGCGCTTCG 600
Db 541 GAGGAGGCTTTCACGAGCACCAAGTGCCTGCGCGCGCGCGCTTGTGACAACTGCGCTTCG 600
QY 601 GACGCTATGCAAGGTTTACACTGAGGCGCACCGCGCGCGCGCGAGGAGGAGAGAGCCCG 660
Db 601 GACGCTATGCAAGGTTTACACTGAGGCGCACCGCGCGCGCGCGAGGAGGAGAGAGCCCG 660
QY 661 CTGACGGAAGACCTACAAGAGATCAACTGCGTGAAGGCTGCGCTTATCGCGCGCGAC 720
Db 661 CTGACGGAAGACCTACAAGAGATCAACTGCGTGAAGGCTGCGCTTATCGCGCGCGAC 720


```

721 AAGCTGGTGAAGTGTGTCGCCCACTACGCGACCGAGATCGCTGCCGATCGCGCGGGGT 780
|||||
721 AAGCTGGTGAAGTGTGTCGCCCACTACGCGACCGAGATCGCTGCCGATCGCGCGGGGT 780
|||||
781 GTGAGCTGGACACCGTATCCGCGCAAGGCGCATTTGAGGCGATTTGTAACGGCATGGAC 840
|||||
781 GTGAGCTGGACACCGTATCCGCGCAAGGCGCATTTGAGGCGATTTGTAACGGCATGGAC 840
|||||
841 ATTGAGAGTGGACACCGCAAGCGACGACGATCTCTGTCGCGCTAGCAGCAGACAGC 900
|||||
841 ATTGAGAGTGGACACCGCAAGCGACGACGATCTCTGTCGCGCTAGCAGCAGACAGC 900
|||||
901 GTCTACGCGCGCAAGGCGCGCCCAAGGAGGCGCTGCAAGCGCGAGCTGGGCTGCGCTG 960
|||||
901 GTCTACGCGCGCAAGGCGCGCCCAAGGAGGCGCTGCAAGCGCGAGCTGGGCTGCGCTG 960
|||||
961 GACCCACGCGCGCGCTGTCGCTTCTGCGCTTCTGCGCGCTGAGGAGCAGAGGCTGGAC 1020
|||||
961 GACCCACGCGCGCGCTGTCGCTTCTGCGCTTCTGCGCGCTGAGGAGCAGAGGCTGGAC 1020
|||||
1021 ATCATCTGCGCGCGCTGCGCAAGATCTGCGCCACCGCCCAAGGTGACATCGCCATCTG 1080
|||||
1021 ATCATCTGCGCGCGCTGCGCAAGATCTGCGCCACCGCCCAAGGTGACATCGCCATCTG 1080
|||||
1081 GGTACGCGCAAGGCGCGCTGACGAGGCTGTGACGCGCATCGGCAACGATACAGGCG 1140
|||||
1081 GGTACGCGCAAGGCGCGCTGACGAGGCTGTGACGCGCATCGGCAACGATACAGGCG 1140
|||||
1141 CGCGCAAGGCGGTGATGATGTCGCGCGCGCTGCGCGCATCTACCGCGCGCGCG 1200
|||||
1141 CGCGCAAGGCGGTGATGATGTCGCGCGCGCTGCGCGCATCTACCGCGCGCGCGCG 1200
|||||
1201 GACTTCATGCTGGTGGCTGCGCTGCGCGCTGCGCGCATCTGCGCGCATCTGCGCGCATG 1260
|||||
1201 GACTTCATGCTGGTGGCTGCGCTGCGCGCTGCGCGCATCTGCGCGCATCTGCGCGCATG 1260
|||||
1261 CACTACGCTACGCTGCGCGTGGTACGCTGCGCGCGCGCGCGCGCGCGCGCGCG 1314
|||||
1261 CACTACGCTACGCTGCGCGTGGTACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1314
|||||

RESULT 7
AF433156 5856 bp DNA linear PLN 23-AUG-2002
LOCUS Chlamydomonas reinhardtii granule-bound starch synthase I (STA2)
DEFINITION gene, complete cds.
ACCESSION AF433156
VERSION AF433156.1 GI:16755882
KEYWORDS Chlamydomonas reinhardtii.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonadaceae; Chlorophyta; Chlorophyceae; Volvocales;
Eukaryota; Viridiplantae; Chlamydomonas.
REFERENCE
AUTHORS Watebled,F., Buleon,A., Bouchet,B., Ral,J.P., Lienard,L.,
Dauville,D., Bunderup,K., Dauville,D., Ball,S. and D'Hulst,C.
TITLE Granule-bound starch synthase I: a major enzyme involved in the
biogenesis of B-crystallites in starch granules
JOURNAL Eur. J. Biochem. 269 (15), 3810-3820 (2002)
PUBMED 12153578
REFERENCE
AUTHORS Watebled,F., Ball,S.G. and D'Hulst,C.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-2001) UMR8576 du CNRS, Laboratoire de Chimie
Biologique, Université des Sciences et Technologies de Lille,
Villeneuve d'Ascq CEDEX 59655, France
FEATURES
source
location/Qualifiers
1..5856
/organism="Chlamydomonas reinhardtii"
/db_xref="taxon:3055"
1343..>4911
/gene="STA2"
join(1343..1781,2151..2340,2484..2618,2815..2973,
```

```

3186..3373,3582..4117,4295..4911)
/gene="STA2"
/product="granule-bound starch synthase I"
join(1480..1781,2151..2340,2484..2618,2815..2973,
3186..3373,3582..4117,4295..4911)
/gene="STA2"
/note="similar to the sequence deposited in GenBank
Accession Number AF026420; GBSSI"
/codon_start=1
/product="granule-bound starch synthase I"
/protein_id="AAL28128.1"
/db_xref="GI:16755883"
/translation="MAVASRSPRSARPIVINAASFVKKTKANOLLRELARGSARKST
SRSAVTGATGATCADIIVMAAEVAPWSKTGDLGDTGGLEIELVKGRHVMITAPRY
DYADAWDTISVVDMIGEHKVFHSIKKGVHRVWDHPFLAKVMGKTGSKLYPRSG
ATYLDNHRKALFCKAALEAAEAVLPFGEDCVFANDWSALVPLLLKDEYQPKGF
TKASLVLIHNFALFQGWEEAFKDKLPOAAFDKLAFLSDGYAKVYTEATMEDEKP
PLTKTKYKINWLKGIILADKLVTSYENYATEIADAAGGVELDVTIRAKISGLIVN
GMDLEENFKDKFLSVYDQNSVYAGKAAKEALQALGLPVDPTAPLPAFIRGLEE
QKQVDIILAAPLKILATPKVQIILGTGKAAEKLVAIGIKYKGRAGVVKFSAFLA
HMLTAGADEMLVPSRFEPGLIQLHAMHYGTVPVASTGGGLDVTKEGTVGFHMGALN
PKLDEADADALATVRRASEVAGRYPENVANCISQDLWSKPAQKEGLLEEVY
GKGVATAKKEEIKVPVAKIPLDLPVSYAPNTLKPVSAVEGNGAAAPKVGITAPA
MGARATPSPGSPAATPKVTVKYPALPAKPKTAGLKLAGEASTTSTSENGAASN
GNGNGASASKISAAKPLVSAATKSA"
BASE COUNT 1073 a 1816 c 1852 g 1115 t
ORIGIN
Query Match 37.9%; Score 603.8; DB 8; Length 5856;
Best Local Similarity 81.2%; Pred No. 6.6e-58;
Matches 795; Conservative 0; Mismatches 7; Indels 177; Gaps 1;
QY 792 CACCGTCTCCGCGCAAGGCGCATTTGAGGCGCATTTGTGAACGGCGATGGACATGAGGAGTG 851
Db 3570 CTCCTCCGCGCGCAAGGCGCATTTGAGGCGCATTTGTGAACGGCGATGGACATGAGGAGTG 3629
QY 852 GAACCCCAAGACCGACAAAGTTCTCTGTCGCGCTACGACCGACGACGCTCTACCGCG 911
Db 3630 GAACCCCAAGACCGACAAAGTTCTCTGTCGCGCTACGACCGACGACGCTCTACCGCG 3689
QY 912 CAAGCGCGCGCGCAAGGAGCGCTGCGAGCGCGAGCTGGCGCTGCGCTGTGACCGCCCGC 971
Db 3690 CAAGCGCGCGCGCAAGGAGCGCTGCGAGCGCGAGCTGGCGCTGCGCTGTGACCGCCCGC 3749
QY 972 CCCCCTGTTGCGCTTTCATCGCGCGCTGAGGAGCGACGAGGCTGTGACATCATCTCTGCG 1031
Db 3750 CCCCCTGTTGCGCTTTCATCGCGCGCTGAGGAGCGACGAGGCTGTGACATCATCTCTGCG 3809
QY 1032 CGCCCTGCGCGCATCTGCGCCACCGCCCAAGGTGCGAGATCGCCATCTGCGTACCGGCAA 1091
Db 3810 CGCCCTGCGCGCATCTGCGCCACCGCCCAAGGTGCGAGATCGCCATCTGCGTACCGGCAA 3869
QY 1092 GCGCGCTACGAGAACTGTGTGAACGCCATCGGCGACCAAGTACAAAGGCGCGCGCGCAAGG 1151
Db 3870 GCGCGCTACGAGAACTGTGTGAACGCCATCGGCGACCAAGTACAAAGGCGCGCGCGCAAGG 3929
QY 1152 CGTGTGCAAGTTCTCGCGCGCGCTGCGCGCACATGCTCACCGCGCGCGCGCGACTTCAGCT 1211
Db 3930 CGTGTGCAAGTTCTCGCGCGCGCTGCGCGCACATGCTCACCGCGCGCGCGCGACTTCAGCT 3989
QY 1212 GGTGCGCTGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCT 1271
Db 3990 GGTGCGCTGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCT 4049
QY 1272 CGTGTGCGGTGAGTACCTCCACCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCG 1331
Db 4050 CGTGTGCGGTGAGTACCTCCACCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCG 4109
QY 1332 CTTCCACA----- 1339
Db 4110 CTTCCACAGTACGCGCGCGCTTCGCGTGGTGGTGAACCTTGTCCTGAGTGGTGGT 4169
QY 1340 ----- 1339
```


Db 4170 GCGGGGCGCTTGACGGGCTCTCTGACGCTGCTACCCCTCTGCTGCGCGCGCCAGGAACCC 4229
 QY 1340 ----- 1339
 Db 4230 ACGGTATCGGAGACCGCTGGCGCATGCTGACGGCTCCCTTTTCCCTCTGTCCCTGAT 4289
 QY 1340 -----TGGGCGCGCTGAACCCGACAAAGCTGGACGAGCTGACCGGAGCGCCCTGCGCGC 1394
 Db 4290 CGCAGTGGCGCGCTGAAACCCGACAAAGCTGGACGAGCTGACGGCGACGCCCTGGCGCG 4349
 QY 1395 CACCGTGGCGCGCTGACGAGGCTGTTTGGGGCGCGCGCTACCCCGAGATGGTGGCCAA 1454
 Db 4350 CACCGTGGCGCGCTGACGAGGCTGTTTGGGGCGCGCGCTACCCCGAGATGGTGGCCAA 4409
 QY 1455 CTTGCATCAGCCAGGACCTGCTGCTGCTCAAGCCGCGCCAGAGTGGAGGGCGCTGTGA 1514
 Db 4410 CTGCATCAGCCAGGACCTGCTGCTGCTCAAGCCGCGCCAGAGTGGAGGGCGCTGTGA 4469
 QY 1515 GGAGTGTGTATCGGCAAGCGCGCGCTGGCCACCGCCAAAGAGGAGGATCAAGTGCC 1574
 Db 4470 GGAGTGTGTATCGGCAAGCGCGCGCTGGCCACCGCCAAAGAGGAGGATCAAGTGCC 4529
 QY 1575 CGTTCGCCAGAGATCCCC 1593
 Db 4530 CGTTCGCCAGAGATCCCC 4548

RESULT 8
 AF286320 2289 bp DNA linear PLN 30-OCT-2000
 LOCUS Triticum aestivum granule bound starch synthase I (GbsSI) gene,
 DEFINITION complete cds.
 ACCESSION AF286320
 VERSION AF286320.1 GI:11037535
 KEYWORDS
 SOURCE Triticum aestivum.
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.
 1 (bases 1 to 2289)
 McCue,K.F., Hukman,W.J., Tanaka,C.K. and Anderson,O.D.
 Starch Branching Enzymes Sbel and Sbe2 from Wheat (Triticum
 aestivum cv. Cheyenne): Molecular Characterization, Developmental
 Expression, and Homolog Assignment by Differential PCR
 Unpublished
 2 (bases 1 to 2289)
 McCue,K.F. and Anderson,O.D.
 Direct Submission
 Submitted (11-JUL-2000) United States Department of Agriculture,
 Agricultural Research Service, 800 Buchanan Street, Albany, CA
 94710-1105, USA

FEATURES
 source location/Qualifiers
 1..2289
 /organism="Triticum aestivum"
 /cultivar="Cheyenne"
 /db_xref="taxon:4565"
 /tissue_type="endosperm"
 204..2021
 /gene="GbsSI"
 204..2021
 /gene="GbsSI"
 /EC_number="2.4.1.21"
 /note="glucosyltransferase; ADP glucose glucosyl
 transferase; waxy; Wx"
 /codon_start=1
 /product="granule bound starch synthase I"
 /protein_id="AAG27624.1"
 /db_xref="GI:11037536"
 /translation="MAALVTSQATSGTILGTRFRAGFCQVRPRSPADAPLGMRT
 TGAAPAKQOSRRKHRGTRCLSMVYRATSGAGMLVFVCAEMAPSKTGLGDLVGG
 LPPMAANGHRVMVISPRVDQYKDAWDISVSEIKVADEYERYFHCYKRGVDRVEV

DHPCFLEKVRGKTKKIYGPDACTDYEDNOLRFSLLCOAALAPRIILNNPNFESGP
 YGEDVYFVNCNDWHITGLLACYLKSNSYSSGIYRTAKVAFCHLNISYOGSRFSDPAQLN
 LPDFKSSFDIDGYDKPVGBKINMKAGIILQADKVLVPSYPYABELLSGERGCEL
 DNIIMLTGITGVINGMDVSEWDPAKDFLAANDIVDTALEGRKALNEALQAEVGLPVD
 RKVPLVAFIGRLEEQKPDVMTAAIPEILKEEDVOIVLLGTGKKFERILKLSVEERFP
 SKVRVVRFNAPLAHOMAGADVLAITSREPCGLIQLQGMRYGTPTCASTGGLVDT
 IMEGKTGFHMHLSVDCNVVEPADVKVVTTLKRAVKVGTTPAYHEMWKNCMIQDLISW
 KGPANKEWDVLLLELGVGSEPGVIGIEIAPLAWENVAAP"

BASE COUNT 480 a 684 c 735 g 390 t
 ORIGIN

Query Match 33.6%; Score 535.8; DB 8; Length 2289;
 Best Local Similarity 61.8%; Pred. No. 2.5e-50;
 Matches 1008; Conservative 0; Mismatches 527; Indels 96; Gaps 6;

QY 5 TGGACATCGTGTGTTGCTGAGTCCGCCCTTGGTCCAAAGAGGGGGCGCTGGCG 64
 Db 436 TGNACCTCGTGTCTGCGCGCGGAGATGGCCCTTGGAGCAAGACCGCGGCTCGCG 495
 QY 65 ATGTGACTGGTGGCCCTGCTTATTTAGCTGGTCAAGCGCGCCCGCTCATGACCATG 124
 Db 496 ACCTCTCGGGGCGCTCCCGCCAGCCATGSCCGCAACGTCACCGGGTCATGTCATCT 555
 QY 125 CCCCTCGCTACGACAGTACGCTGACCGCTGGGACACCTGGTGGTGGAGCATCATGG 184
 Db 556 CCGCGCGCTACGACAGTACGAGGACCGCTGGGACACCGCTGCTCCGAGATCAAGG 615
 QY 185 -----GCGAGAAGGTCGCTACTTCCACTCCATCAAGAGGGGGTCCACCGCG 232
 Db 616 TCGCGACAGATACGAGAGGGTGTAGGTACTTCCACTGCTACAAAGCGCGGGTGGACCGG 675
 QY 233 TGTGTAATCACCACCCCTGTTCTTGGCCAAAGTCTGGGGCAAGACCGCTCCCAAGCTGT 292
 Db 676 TGTTCGTGACACCCCGTCTCTTCTGAGAGGTCTCGGGCAGACCAAGGAGAAATCT 735
 QY 293 ACGGCCCCCGCTCGGCGCTGACTTACTTGACAAACCAAGCGCTTGGCCCTGTCTGCA 352
 Db 736 ACGGCGCGGATCGCGCACCGGACTACGAGGACCAACAGCTACGCTTCAGCCTGCTCTGCC 795
 QY 353 AGGCGCTATTGAGGCTGCCCGCTGCTGCCCTTC-----G 388
 Db 796 AGGCGCGCTTGGGACCCAGATCCTCGACTCAACAAACCCATACTTCTCGGAC 855
 QY 389 GCGCGCGGAGGACTGCTCTTCTGGCCAAAGTCTGGCACTCCGCCCTGGTGGCGCTCC 448
 Db 856 CCTACGGGGAAGACGTGTGTTCTGTCACAGACTGGCACACGGGCTCTTGGCCTGCT 915
 QY 449 TGTGAAGGACGAGTACACAGCCCAAGGGCGAGTTACCAAGGCCAAGTCGCTGGTGA 508
 Db 916 ACCTCAAGACCAACTACCACTGTCAGTCCAGTCCATATAGGACGCCAAGTAGCGTTTGA 975
 QY 509 TCCCAACATCGCTTCCAGGGCGCATGTGGAGAGGCTTTCAGGACGACAGAGCTGC 568
 Db 976 TCCCAACATCTGATCAGGCGCGCTTCTCTTTCAGGACTTTCGCGAGCTCAACCTGC 1035
 QY 569 CCGAGCGCGCTTTGACAAGCTGGCTTCTCGAGCGGCTATGCCAAGGTTTACACTGAGG 628
 Db 1036 CCGACAGGTTCAAGTCTGCTTCTGACTTCAUGACGGCTACGACAG-----1082
 QY 629 CACCCCGCATGAGGAGGACGAGAGAGCCCGCTGACGGGAAGACCTACAGAGATCA 688
 Db 1083 -----CCGCTGGAGGGCGC-----AAGATCA 1104
 QY 689 ACTGCTGAAGGTGGCATTAATGCGCGCGGACAGCTGGTACTGTGTCGCCCAACTACG 748
 Db 1105 ACTGATGAAGCGCGGATCTCTGACGCGCCACAAGGTGCTCAGCTGAGCCCTACTACG 1164
 QY 749 CGACCGAGATCGCTGCCGATGCTCCCGCGGTGTGGAGCTGGACACCGCTCATCGCGCA 808
 Db 1165 CGGAGGAGCTCATCTCGCGCAAGCCAGGGCTCGAGCTCGACACATCATGCGCTCA 1224
 QY 809 AGGCATGAGGCGATTGTGAACCGCATGAGACATTGAGGAGTGGAAACCCCAAGACCGACA 868

Db	1225	CGGGCATACCGGCAATCGTCAACGGCATGGAAGCTACGGAGTGGGAGCCCGCCCAAGGACA	1284
QY	869	AGTTCCTGCTCGCGCCCTACGACCAAGAAAGCGTCTACGCGGCAAGCGCGCCGCAAGG	928
Db	1285	AGTTCCTCGCGCCCAACTACGAGGTACACCGCGTTGGAGGGGAAGCGCTGAACAAGG	1344
QY	929	AGGCGCTCGAGCGCGAGCTGGGCGCTGTGTGGACCCACCGCCCCCTGTTCGCGCTTCA	988
Db	1345	AGGCGCTCGAGCGCGAGGTGGGCGCTGGCCGCTGGACCGGAAGGTGCCCTGTGGCGCTTCA	1404
QY	989	TGCGCGCGCTGGAGGACGAGAGGGTGTGGACATCATCTTGGCGCGCCCTGCCCAAGATCC	1048
Db	1405	TCGGCAGGCTGGAGGACGAGAGGCGCCGACGTGATGATCGCCGCCATCCGGAGATCT	1464
QY	1049	TGGCCACCCCAAGCTGCGAGATCGCCATCTTGGGTACCGGCAAGCGCGCCCTACGAGAAGC	1108
Db	1465	TGAAGGAGGAGGACGTCCAGATCGTCTCTTGGGCACCGGAAGAAGTTTGGCGGC	1524
QY	1109	TGTTGAAGGCCATPGGCACCAAGTACAAGGGCGCGCCCAAGGGCGTGGTCAAGTTCTCGG	1168
Db	1525	TGCTCAAGAGCGTGGAGGAGAAGTTCCCGAGCAAGGTGAGGCGCGTGGTCAGGTTCAACG	1584
QY	1169	CGCCCTGGCGACATGCTCACCGCGCGCGGCGACTTCATGCTGTTGCCCTCGCGCTTCG	1228
Db	1585	CGCGCTGGCTCACAGATGATGGCGCGCGCGACGTGCTTGGCGTCAACAGCGCCTTCG	1644
QY	1229	AGCCCTCGCGCCTGATCCAGCTGCACGCCATGCATCTAGCGTACCGTGCCCGTGGTAGCCT	1288
Db	1645	AGCCCTCGCGCCTCATCCAGCTCCAGGGGATGCGTACGGAACGCCGTGCGCGTGC	1704
QY	1289	CCAGCGCGCGCTGTGACACCGTCAAGGAGGGCGTCAACCGGTTCCACATGGGCGCCC	1348
Db	1705	CCACCGCGGGCTCTGACACGATCATGGAGGGCAAGACCGGGTTCACATGGCGCCACC	1764
QY	1349	TGA-----ACCCCGCAAGCTGGAGAGGCTGACGCCGACGCCCTGGCCCGCACCGTGC	1402
Db	1765	TCAGCTCGACTGCACAGTGTGTGGACCGCGCGACGTGNAGAGTGTGTGACCCCTGA	1824
QY	1403	GCCGTGCCAGCGAGGTGTTTGGCGGCGCGCGTACCCCGAGATGTTGGCCCAACTGCATCA	1462
Db	1825	AGCGCGCGTCAAGGTCGTGGCACGCCAGCCCTACCATGAGATGTTCAAGAAGCTGCATGA	1884
QY	1463	GCCAGGACCTGTCTGTGTTCCAAGCCCGCCAGAGTGGGAGGGCCCTGCTGGAGGAGTGG	1522
Db	1885	TCCAGATATCTCTCTGGAAGGGGCCAGCAAGACTTGGGAGACGTGTCTTGGAACTGG	1944
QY	1523	TGTACGCAAGGGCGCGCTGGCCACCGCCCAAGAAGGAGGATCAAGTTCGCCGTTCGCG	1582
Db	1945	GG---GTGAGGGGAGCGACCGAGGGGTCTATCGCGAGGAGATTCGCGCGCTCGCCATGG	2001
QY	1583	AGAAATCCCC	1593
Db	2002	AGAACTCGC	2012

RESULT 9	
AF486514	
LOCUS	1812 bp mRNA linear PLN 02-JUL-2002
DEFINITION	Hordeum vulgare cultivar oderbrucker granule bound starch synthase I mRNA, complete cds.
ACCESSION	AF486514
VERSION	AF486514.1 GI:21667427
KEYWORDS	.
SOURCE	Hordeum vulgare.
ORGANISM	Hordeum vulgare.
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
	Pooideae; Triticeae; Hordeum.
REFERENCE	1 (bases 1 to 1812)
AUTHORS	Patron,N., Smith,A., Fahy,B., Hylton,C., Maldrett,M., Rossnagel,B. and Denyer,K.
TITLE	A mutation in the 5' non-coding region of the barley GBSSI gene alters its temporal and spatial expression and reduces GBSSI

activity and amylose content in the endosperm
Unpublished
2 (bases 1 to 1812)
Patron, N.
Direct Submission
Submitted (22-FEB-2002) Metabolic Biology, John Innes Centre,
Norwich Science Park, Norwich, Norfolk NR4 7UH, UK
Location/Qualifiers
1. .1812
/organism="Hordeum vulgare"
/culturivar="oderbrucker"
/db_xref="taxon:4513"
1. .1812
/note="alpha glycosyl transferase"
/codon_start=1
/product="granule bound starch synthase I"
/protein_id="AAM74048.1"
/db_xref="GI:21667428"
/translation="MAALASQLATSGTLVGTDRFRPFGOGLRPNRPADAALGMRT
IGASAPQSRKHAHRSRCLSVVSATSGMNLVFGAEMAPWSKTGGLDVGLGLP
PMAANGHRVWVSPPRQDKADWDISVISEIKVADEYERFEHCYKRGVDRVFIDH
PWLEKYRGTKIKYIDQAGTDYEDNQRFSLKIQQAALAPRLNLNNPYPSPGPG
EDVFVCDNHDHTGLLACYLKSNTQSNIGITAVAFICVHNSYOGRESFDDFAQLNLP
DRFKSVGTDIDGYDKPVEGRKINWKGILQAOKVLTVSPYAEELISGARCELDN
IMRLTGITGVNGMDVSEWPTDKDLFVANDITITALEAKALKEALQAQEVLPVDRK
VPLVATGIFEEQGDPMVMTAIPILKEEDVOITLLIGTKKKFKLLKSGWEKFPQK
VRVVRNAPLAFHOMAGADLLATSRFPCGLIOLQGMRYGTPCYCASTNGGLVDITV
EGKTFHMGRLSVDCNVVEADVAKKVATTLKRAVYVGTTPAYQEMVKNCHIQLDSLWKG
PAKNNEVDVLELGVGSEPGVEETPLAMENVAAP"

363 a 564 c 597 g 288 t

BASE COUNT
ORIGIN

Query Match	33.3%;	Score 530;	DB 8;	Length 1812;	
Best Local Similarity	61.6%;	Prod. No. 1.2e-49;			
Matches	985;	Conservative	0;	Mismatches 520;	Indels 93; Gaps 5;

QY	5	TGCACATCGTGATGGTTCGCTGAGGTGCGCCCTTGGTCCAAAGACGGCGCCCTGGCGG 64
DB	227	TGAACCTCGTGTTCTGCGGCGGAGATGGCGCCCTGGAGCAAGACCGCGGCCCTTGGCG 286
QY	65	ATGTGACTGGTGGCCCTGGCTATTGAGCTGGTCAAGCGGGGCCACCGCGGTATGACCATTTG 124
DB	287	ATGTCTCGGGGCGCTCCCGCGGCATGGCGGCCAACGGTCAACGGGTGATGGTGGTCT 346
QY	125	CCCTCGCTACGACGACGATGAGCTGAGCGCTGGGACACCTCGGTGCTGGACATCATGG 184
DB	347	CCCCGCGCTACGATCAGTACAAGGACGCGTGGGACACCAAGCGCTATCTCCGGAGATCAAGG 406
QY	185	-----GCAGAAAGTCCGCTACTTCCACTCCATCAAGAAAGGCGGTGCACCGCG 232
DB	407	TGCTGACGAGTACGAGAGGTGAGGTTCCTCCACTGTACAAAGCGCGGAGTGGACCGCG 466
QY	233	TGTGGATTGACCAACCCCTGGTTCCTGGCCCAAGGTCTGGGGCAAGACCGGCTCCAGCTGT 292
DB	467	TGTTATCAGCAACACCGTGTCTCTGGAGAAGGTCTCGGGGCAAGACCAAGSAGAATCT 526
QY	293	AGGGCCCCGCTCGGCGCTGACTACCTGAGCAACCAAGCGCTTCGCGCTGTCTGCA 352
DB	527	ACGGGCCCGACCGCGCAGACTATGAGGACAACCAAGCGCTTCAGCCTTCCTGCC 586
QY	353	AGGCCGCTATTAGGCTGCCCGCTGCTGCCCTTC-----G 388
DB	587	AGGCAGCACTCGAGGCACCCAGGATCTCAACCTCAACAACACCCCTACTTTCTGGTC 646
QY	389	GGCCCGGCGAGGACTCGCTTCGTGGGCCAACGACTGGCACTCGGCCCTGGTGGCCGTC 448
DB	647	CTTACGGGGAAGACGTGGTGTCTGTGTGAACGACCTGGCACAACCGGCCCTTCGGCGCTGT 706
QY	449	TGCTGAAGGACGAGTACCAAGCCCCAAGGCGCAGTTCAACAAAGGCAAGTGGTCTGGCTA 508
DB	707	ACCTCAAGCAACTACAGTCCAAATGGCATCTACAGGACGGCCAGGTGGCTTCTGCA 766
QY	509	TCACAAACATCGCCTTCCAGGGCCCGCATGTGGGAGGAGGCTTTTCAAGGACACGAAGCTGC 568

```

|||||
767 TCACAAACATCTCGTACAGGCGCGTCTCTCTTCGACGACATTGCGCAGCTCAACCTGC 826
QY CCCAGGCGCCCTTTGACAAAGCTGGCTTCGCGAGCGGTGTAAGCAAGTTTACACTGAGG 628
Db CCGACAGGTTCAAGTCTGCTCTTCTGACTTATTCAGCGCTACGACAAG----- 873
QY CCACCCCCATGGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 688
Db CCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 895
QY ACTGGCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 748
Db ACTGGATGAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 955
QY CGACCCAGAGATCGCTGCGGATGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 808
Db CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1015
QY AGGGCAATCAGGCGCATGTTGAAGCGCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 868
Db CGGGATCAGCGGATGTTGAAGCGCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1075
QY AGTCTCTGCTGCGGCTTACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 928
Db AGTCTCTGCGCTCAACTAGGACATCAACACCGGCTTGAGGCGGAGGAGGAGGAGGAGGAGG 1135
QY AGGCGCTGACGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 988
Db AGGCGCTGACGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1195
QY TCGGCGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1048
Db TCGGCGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1255
QY TGGCCACCCCAAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1108
Db TGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1315
QY TGGTCAAGCGCATCGGACCAAGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1168
Db TGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1375
QY CGCCCGCTGGCGCACATGCTCACCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1228
Db CGCCCGCTAGCTCACAGATGATGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1435
QY AGCCCTGGCGCTGATCCAGGTGACGCGCATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1288
Db AGCCCTGGCGCTGATCCAGGTGACGCGCATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1495
QY CCACCGCGCGCTGCTGACACCGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1348
Db CCACCGCGCGCTGCTGACACCGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1555
QY TGA-----ACCCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1402
Db TCAGCGTGCAGTCAACGCTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1615
QY GCGGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1462
Db AGCGGGCGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1675
QY GCCAGGAGCTCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1522
Db TCCAGGAGTCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1735
QY TGTACGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
Db GGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1773

```

RESULT 10

HWAXYR
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

HWAXYR
 Barley mRNA pcwx27 for waxy locus.
 X07932
 X07932.1 GI:19128
 starch synthase; waxy locus.
 Hordeum vulgare subsp. vulgare.
 Hordeum vulgare subsp. vulgare.
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Hordeum.
 1 (bases 1 to 2311)
 Zohde, W.
 Direct Submission
 Submitted (21-JUN-1988) Zohde W., Max Planck Institute, Erwin Baur
 Institut, D-5000 Koln 30, FRG
 2 (bases 1 to 2311)
 Rohde, W., Becker, D. and Salamini, F.
 Structural analysis of the waxy locus from Hordeum vulgare
 Nucleic Acids Res. 16 (14B), 7185-7186 (1988)
 88303345
 2970062
 PUBMED

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED

COMMENT
 FEATURES
 source

source

CDS

/organism="Hordeum vulgare subsp. vulgare"
 /strain="Vogelsanger Gold"
 /db_xref="taxon:112509"
 /map="chromosome 1"
 /clone="pcwx27"
 /clone_lib="EMBL4"
 226..2037
 /note="starch synthase"
 /codon_start=1
 /protein_id="CAA30756.1"
 /db_xref="GI:19128"
 /db_xref="SWISS-PROT:P09842"
 /translation="MAALATSLATGTVLGVTDREPRFGQLRPNPADAALGMRT
 IGASAAKQSRKAHRGSRCLSVVVSATGSGNNLVFGAEAPWSTGGLGVLGGLP
 PAMRANGHRVWVSPRYDQKADWTISLSEIKVADEYERVFHYKRGVDRFIDH
 PWFLEKVRGTEKTYGPDAGTIEDNQORFSLSCQAALAPRLINNNYFSGPYG
 EDVFEVNDWHITGLLACYLKNYSNGYRTAKVAFCHINISYGRFDFDPAQNLNLP
 DRKSSDFIDGYDKEVGRKINWKAQILQADKLVITSPYAEELISGEARGCLDN
 IMRLTGITGVNGMDVSEWPTKDFLAVNYDITLLEAKALKEALQAEVGLPVDK
 VPLVAFIGRLPEQGPDMVIAIPILKEEDVQIILCTGKKKFKLLKSMEEKPPGK
 VRVVRNAPLAHOMAGADLLAVTSREPEGLIQLQMRVGTGTCVCASTGGGLVDITV
 EGKTFHMGRLSDVNCNVVPADVKKVAILLRAVKVVGTPAYQEMVKMCMQDLSWKG
 PAKNEDVLLLELGVGSEPGIVGEEIAPLAVENVAAP"
 2276

polyA_site

polyA_site

polyA_site

polyA_site

polyA_site

polyA_site

polyA_site

polyA_site

polyA_site

polyA_site

polyA_site

polyA_site

polyA_site

polyA_site

polyA_site

polyA_site

polyA_site

Query Match 33.3%; Score 530; DB 8; Length 2311;
 Best Local Similarity 61.6%; Pred. No. 1.1e-49;
 Matches 985; Conservative 0; Mismatches 520; Indels 93; Gaps 5;

QY 5 TGGACATCGTGATGGTTCGTCTGCTAGGTCGCGCCCTTGGTCCAAAGACGGCGCGCTGGCGG 64
 Db 452 TGAACCTCGTGTCTGTCGCGCGCGGAGATGGCGCCCTGGAGCAAGACCGCGCGCTGGCG 511
 QY 65 ATGTGACTGTGGCGCTGCTTATGTAGCTGGTGAAGCGGCGGCGGCGCTCATGACCATG 124
 Db 512 ATGTCTCGGGGCGCTCCCGCGGCGCATGGCGGCGGCGGCGGCGGCGGCTGCTGCT 571
 QY 125 CCCCTCGCTACGACGACGACGCTGACGCGCTGGGACACCTCGGCTGCTGACATCATGG 184
 Db 572 CCCCGCTACGATGATACAGGACGCTGGGACACCGGCGGCGGCGGCGGCGGCGGCGG 631

```
QY 185 -----GCGAAGGTCCGCTACTTCCACTCCATCAAGAAGGCGGTGCACCGCG 232
Db 632 TCGCTGAGGAGTACGAGAGGTGAGGTCTTCCACTGCTCAAGCGCGGAGTGGACCGCG 691
QY 233 TGTGATTGACACCCCTGTTCTCTGCGCAAGTCTGTGGGCAAGACCGGCTCCAAAGTGT 292
Db 692 TGTTCATCGACACCCGTTGTTCTGTGAGAAGTTCGCGGCAAGACCAAGGAGAAGATCT 751
QY 293 ACGGCCCCGCTCGGCGCTGACTACTCTGGACACACCAAGCGCTTCGCCCTGTCTCTGCA 352
Db 752 ACGGCCCCGACGCGCGCACACTATGAGGACCAACACGACGCTTTCAGCCCTTCTCTGCC 811
QY 353 AGGCGCTATTGAGGCTCCCGCGTGTGCGCTTC-----G 388
Db 812 AGGAGCACTCGAGCACCCAGGATCTCACTCAACAACAACCCCTACTTTCGTGTC 871
QY 389 GCCCGGCGGAGACGCTCTCTCTGTCGCAAGCTGCACTCGCCCTGTCGCGCGTCC 448
Db 872 CCTACGGGGAAGACGTGTGTGTCGTCGCAAGCTGCACTGCGCGCTCTCTGCGCTGCT 931
QY 449 TGCTGAAGSAGAGTACCAGCCCAAGGCGCAGTTTCAACAAGGCCCAAGTCTGCTGCTG 508
Db 932 ACCTCAAGAGCAACTTACCAGTCCATGTCATCTACGACGCGCCCAAGTGTGCTTCTGCA 991
QY 509 TCCACAACATCGCCTTCCAGGCGCGCATGTGGGAGGAGGCTTTCAGGACACAGAACTGC 568
Db 992 TCCACAACATCTCGTACCAGGCGCGCTTCTCTCTCGAGGACTTTCGCGAGCTCAACCTGC 1051
QY 569 CCCAGCGCGCTTTCAGCAAGCTGGCTTCTCGGACGCGCTATGCAAGGTTTACACTGAGG 628
Db 1052 CCGACAGGTTCAAGTCTGCTCTGCACTTCAATTGACGGCTTACGACAAAG----- 1098
QY 629 CCACCCCATGAGGAGGACGAGAGAGCCCGCTGACGGAAGACCTTACAAGAAGATCA 688
Db 1099 -----CCCGTGGAGGGGCGC-----AGATCA 1120
QY 689 ACTGGCTGAAGGTTGGCATATTCGCGCGCGACAAAGCTGGTGTGCTGCGCCCACTAG 748
Db 1121 ACTGGATGAAGCGCGGATCTCTGACGCGCGCAAGGTGCTGAGCTGAGCGCCCTACTACG 1180
QY 749 CGACCGAGATCGCTGCGCATGCGCGCGGGGTGTGGAGTGTGACACCTGCATCCGCGCA 808
Db 1181 CTGAGGAGCTCATCTCTGCGGAAGCGAGGCGCTGCGAGCTGCAACATATCGCGCTCA 1240
QY 809 AGGCACTGAGGCGCATGTGAAGCGCATGGACATTGAGGAGTGGAAACCCCAAGACCGACA 868
Db 1241 CGGGATACCGGAATCGTCAACGCGCATGGAGCTGAGTGTGAGTGGACCCCAAGGACCA 1300
QY 869 AGTTCTGTCTGCGCCCTACGACCAAGACAGGCTCTACCCGCGCAAGGCGCGCGCAAGG 928
Db 1301 AGTTCTGCGCGTCACTACGACATCACCAACCGCGTTGAGGCGCAAGGCGCTGAACAAGG 1360
QY 929 AGGCGCTGAGCGGAGTGGCGCTGTCGTCGACCCCAAGCGCGCGCTGTCGCGCTTCA 988
Db 1361 AGGCGCTGAGCGGAGTGGCGCTGTCGTCGACAGGAGTGGCGCTGTCGCGCTTCA 1420
QY 989 TCGGCGCGCTGAGGAGCAAGAGGTGTGACATATCTGCGCGCGCGCGCCCAAGATCC 1048
Db 1421 TCGGAGGCTGAGGAGCAAGAGGCGCGCGCGCTGATGATGCGCGCGCATCCCGGAGATCC 1480
QY 1049 TGGCCACCCCAAGGTGCGCATCGGCATCTGGGTACCGGCAAGCGCGCGCTACGAGAAGC 1108
Db 1481 TGAAGAGGAGGAGCGTCCAGATCATCTCTTGGCACCGCGGAAGAAGTITGAGAAGC 1540
QY 1109 TGGTGAAGCGCATCGCACCAAGTACAAGGCGCGCGCGCGCGCGCGTGTGCTGCTCGG 1168
Db 1541 TGTCTAAGAGCATGAGGAGGAAGTTCCTCCCGGCAAGGTGAGGCGCGGTGCTGAGTTC 1600
QY 1169 CGCCCTGCGCGCATGCTCAACCGCGCGCGCGCGCGCTTCACTGTGTCGCTGCGCGCTG 1228
Db 1601 CGCCGCTAGCTCACCAGATGATGCGCGCGCGCGCGCGCTTCTGCTGCTGTCACCAAGCGCTG 1660
```

```
QY 1229 AGCCCTGCGGCGTATCCAGCTGACAGCGCATGCACCTACGTTACCGTCCCGCTGGTAGCCT 1288
Db 1661 AGCCCTGCGGCGTATCCAGCTGACAGCGCATGCACCTACGTTACCGTCCCGCTGGTAGCCT 1720
QY 1289 CCACGCGGCGCTGTCGACACCGCTCAAGGAGGCGCTCACCGGCTTCCACATGGCGGCC 1348
Db 1721 CCACGCGGCGCTGTCGACACCGCTGTCGAGGCGCAAGACCGGCTTCCACATGGCGGCC 1780
QY 1349 TGA-----ACCCCGACAAAGCTGGACAGGCTGACGCGCACCGCCCTGCGCGCGCCGTCG 1402
Db 1781 TCAGCGTCTGCTGCAACGCTGTCGAGCGCGCGGCTGAAGAAGCTGGCGACACCCCTGA 1840
QY 1403 GCGGTGCGACGAGTGTTCGCGGCGCGCTACCCGAGAGTGTGGCCACTGCATCA 1462
Db 1841 AGCGGCGGCTCAAGTCTGCGGCGCGCTTACGAGGATGTCAAGAAGTTCATGA 1900
QY 1463 GCCAGGACTGTCCTGTCCTCAAGCGCGCGCAAGTGGAGGCGCTGCTGGAGGAGTGG 1522
Db 1901 TCCAGATCTCTCTGGAAGGACCTGCAAGAAGTGGAGGAGCTGCTTCTGGAAGTGG 1960
QY 1523 TGTACGCGCAAGGCGCGCTGGCCACCGCGCAAGAAGAG 1560
Db 1961 GGTGGAGGAGCGAGCGCGGATCGTCGCGGAGGAG 1998

RESULT 11
AF486515 1827 bp mRNA linear PLN 02-JUL-2002
LOCUS Hordeum vulgare cultivar waxy oderbrucker granule bound starch
DEFINITION synphase I mRNA, complete cds.
ACCESSION AF486515
VERSION AF486515.1 GI:21667429
KEYWORDS Hordeum vulgare.
SOURCE Hordeum vulgare
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 1827)
AUTHORS Patron,N., Smith,A., Fahy,B., Hylton,C., Naldrett,M., Rossnagel,B.
and Denyer,K.
TITLE A mutation in the 5' non-coding region of the barley GBSSI gene
alters its temporal and spatial expression and reduces GBSSI
activity and amylose content in the endosperm
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1827)
AUTHORS Patron,N.
JOURNAL Direct Submission
Submitted (22-FEB-2002) Metabolic Biology, John Innes Centre,
Norwich Science Park, Norwich, Norfolk NR4 7UH, UK
FEATURES
source
1..1827
/organism="Hordeum vulgare"
/cultivar="waxy oderbrucker"
/db_xref="taxon:4513"
1..1827
/note="alpha glycosyl transferase"
/codon_start=1
/product="granule bound starch synthase I"
/protein_id="AAM74049.1"
/db_xref="GI:21667430"
/translation="MAAATLSQIA"SGTVLGVDRSAPSMERHAGFOGLRPRKPADGT
FQMRITVGAAPKQSKAHNRCLSVVVRATGSGMNLVFGAEMAPMSKTTGLGVDV
LGLEPPAMAANGHRVMVSPRYDQKADMTSVIETKVADETERVRFPHCKRGVDR
VFIDHPLEKVRGKTKERYIGDADGTEDYDQRFSLCCOALEAPRLINLNINNYF
SPGYGDEVVVCNDWHTGLLACYLAKSNYSNGIYTKAVFCIHNISYQGRFSFDFA
QNLNPDREKSSDFIDGYDKPVEGRKINNMKAGILQADKVLTVSPYIAELISGEARG
CELDNMLRLTGTGIVNGMDSEMDTKDFLAVNYDITTALEAKALKEALQAEVGL
PVDKVPVAFVIGRLEEQGPVMIHAIPELLKEEDYQIILLGTGKFKPKLLKSMEE
KFPKVRVAVRFNAPLAHOMAGADIAITLRFEPGCLIQLOLMRGTGTCVCASTGGL
VPTIVETGTFHMLSLVDCNVPEPADKKVATTLKRVKVGTPAYQEMKNCMIQD
LSWKGPKNWEDVLELGEVGESEPIVGEIAPLAMEVAAAP"
BASE COUNT 364 a 567 c 602 g 294 t
```

ORIGIN

Query Match 33.2%; Score 528.4; DB 8; Length 1827;
 Best Local Similarity 61.6%; Pred. No. 1.7e-49;
 Matches 984; Conservative 0; Mismatches 521; Indels 93; Gaps 5;

QY 5 TGGACATCGTGTGCTGCTGAGTCCGCCCTTGGTCCAGACAGCGGGGCTGGGCG 64
 DB 242 TGNACCTCGTGTGCTGCGCGCGAGATGGCGCTTGGAGCAAGCGGGGCTTGGCG 301

QY 65 ATGTGACTGGTGGCCCTGCTATTGAGCTGTCAAGCGCGGCCACCGCGTCATGACCATG 124
 DB 302 AGCTCTCGGGCGCTTCCACAGCCATGGCCGCCAAGGTCACCGGGTCATGCTGCTCT 361

QY 125 CCCCTCGTACGACAGTACGCTGAGCGCTGGGACACCTCGGTGGTGGACATATGG 184
 DB 362 CCCCGGCTACGATCAGTACAGGACGCTGGGACACCGGCTCATCTCGGAGTCAAGG 421

QY 185 -----CGGAGAAGGTCCTGCTTCCACTCCATCAGAAAGGCGTGCACCGCG 232
 DB 422 TCGCTACGAGTACGAGAGGTGAGTCTTCCACTGCTACAAAGCGGAGTGGACCGG 481

QY 233 TGTGGATTGACCAACCGCTGTCTTGGCCAAAGTCTTGGGCAAGACCGGCTCCAAAGCTGT 292
 DB 482 TGTTCATCGACCAACCGCTGTCTTGGAGAGGTCCGGGCAAGACCAAGGAGAATCT 541

QY 293 ACGGCCCGCGCTCCGCGCTGACTACTGACCAACCAAGCGCTTCGCGCTGTTCGCA 352
 DB 542 ACGGGCCCGACCGCGCAGGACTAGGAGGCAACCAAGCGCTTCAGCGCTTCTGCGC 601

QY 353 AGGCGGCTATTGAGGTGCGCGGCTGCTGCGCTTC -----G 388
 DB 602 AGGACGCGCTCGAGGACCCAGGATCCTCAACTCAACAAACCCCTACTTTTCGGTC 661

QY 389 GCGCGCGGAGGACGCTGCTTGTGGCCAAAGTCTGGGCACTCCGCGCTGGTCCCGTCC 448
 DB 662 CTAAGGGGAAGACGTGTGTGCTGCAAGCTTGGGCAAGCGGCTTCTGGCGTGT 721

QY 449 TGTGAAGGACGATACCAAGCCAGGCGCAGTTCAACAGGCGCAAGTGGTGTGGGTA 508
 DB 722 ACTCAAGAGCACTACCAAGTCCATGAGGATCTACAGGACGCGCAAGTGGCTTCTGA 781

QY 509 TCCCAACATCGCTTCCAGGCGCGCATGTGGAGGAGCTTTCAAGGACACGAGCTCG 568
 DB 782 TCCCAACATCTGTAACAGGCGCGCTTCTCTTGCAGCACTTGGCGAGCTCAACCTCG 841

QY 569 CCCAGCGCGCTTCAAGAGCTGGCGCTTCGAGCGGTATGCCAAGGTTTACACTGAGG 628
 DB 842 CCGACAGGTTCAAGTCTGCTTCGACTTCATTGACGGCTTAACGACAAAG ----- 888

QY 629 CCACCCCATGAGGAGGACGAGAGCCCGCGCTGACGGGAAGACCTACAGAGATCA 588
 DB 889 -----CCGCTGAGGGGCGC -----AAGATCA 910

QY 689 ACTGCTGAAGGTGGCATTAATCGCGCGCACAAAGTGTGACTGTGCGCCCAACTAG 748
 DB 911 ACTGATGAAGCGCGGATCTGCGAGGCGCAAGGCTGCTAGCGTGAAGCGCTACTAG 970

QY 749 CCACGAGATCGCTCGCATGCCGCGCGGGGTGTGAGCTGAGACACCGTCATCCGCGCA 808
 DB 971 CTGAGGAGCTCACTCTGCGCAAGCGAGGGGCTGGGCTGCAACATCATGCGCTCA 1030

QY 809 AGGCGATTGAGGCGCTTGTGAAGCGCATGACATGAGAGTGAACCCCAAGACCGCA 868
 DB 1031 CGGGATCACCAGATCTGACAGCGATGAGGCTCAGTGTGAGGACCGCCCAAGGACA 1090

QY 869 AGTCTCTGTCGCGCTTACGACCAAGACAGGCTTACCGCGCAAGCGCGCGCAAGG 928
 DB 1091 AGTCTCTGCGCTCAACTACGACATCACCACCGCTTGGAGCGCAAGCGCTGACAAAG 1150

QY 929 AGGCGCTGAGGCGCGAGCTGGCGCTGTGGACCCACCGCGCGCGCTGTTCGCTTCA 988
 DB 1151 AGGCACTGAGGCGCGAGGTGGGGCTGCGGCTGAGACAGGAAGTGGCGCTGGCTTCA 1210

QY 989 TCGCCGCTTGGAGGAGCAGAAAGGTGTGGACATCATCTCTGGCGCGCTGCCAAGATCC 1048
 DB 1211 TCGCAGGCTGGAGGACAGAGGCGCGGACGTGATGATCGCGCCATCCCGGAGATCC 1270

QY 1049 TGGCCACCCCAAGGTGAGATGCCATCTCTGGTACCGCAAGCGCGCTTACGAGAAGC 1108
 DB 1271 TGAAGGAGGAGGACGTCCAGATCATCTCTTGGCACCGGGAAGAAGTTTGAGAAGC 1330

QY 1109 TGTGTACGCGCATCGGACCAAGTACAAAGGCGCGCCCAAGGCGTGTCAAGTCTCGG 1168
 DB 1331 TGTCTAAGACATGAGGAGAAAGTTCCCGGGCAAGGTGAGGCGCGTGTTCAGGTTCAACG 1390

QY 1169 CGCCCTTGGCGCACATGCTCACCGCGCGCGCTTCACTGCTGCTGCCCTCGCGCTTCG 1228
 DB 1391 CGCGCTAGCTACCAAGATGATGCGCGCGCGGACCTTCTGCTGTCACCAAGCGCTTCG 1450

QY 1229 AGCCCTGCGGCTGATCCAGCTGACGCCATGACCTACGCTACCGTACCGTGTGCTAGCCT 1288
 DB 1451 AGCCCTGCGGCTCATCCAGCTCAGGGAATGCGCTATGGAAGCGCGTGTGTGCGCT 1510

QY 1289 CCACCGCGGCTGTGCGACACCTCAAGGAGGCGCTCACCGGCTTCCACATGCGCGCC 1348
 DB 1511 CCACCGGCGGCTGTGCGACACCTCAAGGAGGCGCTTCCACATGCGCGCGCT 1570

QY 1349 TGA-----ACCCGCAAGCTGGAGAGGCTGACCGCGACCGCTTGGCGCGCACCGCTGC 1402
 DB 1571 TCAGGCTCGACTGCAAGCTGTGTGGAGCGCGGACGTGAAGAAGTGGCGCACCGCTGA 1630

QY 1403 GCGTGGCAGGAGGTGTGTTGCGGCGCGCGCTACCCCGAGATGGTGGCGCACTGCATCA 1462
 DB 1631 AGCGGCGGCTCAAGGTGTGTCGCGCACCGCGCTACCAAGGAGTGGTCAAGACTGATGA 1690

QY 1463 GCAGGACCTGTCTTGTCCAAAGCGCGCAGAAAGTGGAGGCGCTCTGGAGAGGTGG 1522
 DB 1691 TCAGATCTCTCTGAGGAGGACCTGCCAAGACTGGGAGGAGCTTCTTGGAACTGATGA 1570

QY 1523 TGTACGCGAGGCGCGCTGGCCACCGCGCAAGAGGAG 1560
 DB 1751 GGTGGAGGAGGAGCGCGGATGCTGCGCGAGGAG 1788

RESULT 12

AF486518
 LOCUS
 DEFINITION
 mRNA, complete cds.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Hordeum vulgare.
 Hordeum vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Hordeum.
 REFERENCE
 AUTHORS
 TITLE
 A mutation in the 5' non-coding region of the barley GBSR gene
 alters its temporal and spatial expression and reduces GBSR
 activity and amylose content in the endosperm
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 1. .1827
 /organism="Hordeum vulgare"
 /cultivar="SB 85750"
 /db_xref="taxon:4513"
 1. .1827
 CDS

AUTHORS Patron, N., Smith, A., Fahy, B., Hylton, C., Naldrett, M., Rossmagel, B. and Denyer, K.
TITLE A mutation in the 5' non-coding region of the barley GBSSI gene alters its temporal and spatial expression and reduces GBSSI activity and amylose content in the endosperm
JOURNAL REFERENCE Unpublished
AUTHORS 2 (bases 1 to 1812)
TITLE Patron, N.
JOURNAL Direct Submission
 Submitted (22-FEB-2002) Metabolic Biology, John Innes Centre, Norwich Science Park, Norwich, Norfolk NR4 7UH, UK

FEATURES Location/Qualifiers
 1..1812
 /organism="Hordeum vulgare"
 /cultivar="CDC Alamo"
 /db_xref="taxon:4513"
 1..1812
 /note="alpha glucosyl transferase"
 /codon_start=1
 /product="granule bound starch synthase I"
 /protein_id="AA074052.1"
 /db_xref="GI:21667438"
 /translation="MAALATSLATSGVILGVTDRFRPFQGLRPNDAALGMRT IGASAPKQSRKHSRRCLSVVYATSGMNLVFGAEMAPWSKTGLGLVGLSLP PMAAAGHRVWVSPRYDQYKDAWTSVISELKVADEVRVFEHCYKRGVDRVIDH PWFLEKVRGKTEKIXGPDAGTDYEDNQGRFSLCOALEAPRILNINNNPFIISGPIG EDVVFVNDWHGLACILYSNIGYRTAKVAFCHNISYGRFSDFDFAQNLPL DRFKSDFIVGYDVEVRGKINNMKAGILQADKIVTSPIYAEELISGEARGCELDN IMLRGITGVINGMDVSDWPTKDFLAVNYDITALEAKNLKALQAEVLGPVDRK VPLVAFIGRLKQGPDMVIAAIPILKEEDVQIILGTGKKKLEKLLKSMKEPFGK VAVRVFNAPLADHOMAGADLLAVTSREPCGLIOLOGMRYGTPCVCASTGLVDLIV EGKTFGHRGLSDCNVVEPADVKVATILKRAVNVVTPAQEWMVKNVCVVDLSWKG PAKNEDVLLLELGVSGSEPGIVGEIEIPLAMENVAAP"
 BASE COUNT 359 a 565 c 599 g 289 t
 ORIGIN

Query Match 33.0%; Score 525.2; DB 8; Length 1812;
 Best Local Similarity 61.5%; Pred. No. 3.9e-49;
 Matches 982; Conservative 0; Mismatches 523; Indels 93; Gaps 5;
 QY 5 TGAACATCGTGTGCTGAGGTGCGCCCTTGGTCCAAAGCGGCGGCTGGCG 64
 Db 227 TGAACCTCGTGTGCTGCGCGGAGATGCGCCCTGGAGCAAGACGCGGCTGGCG 286
 QY 65 ATGTGACTGTCGCTTCCTATGACTGGTCAAGCGCGGCCACCGGTCATGACCATG 124
 Db 287 ATGTCTCGCGGCGCTCCCGCGGCATGGCCGCCAACCGTCAACCGGCTGATGGTCT 346
 QY 125 CCCTCGCTACGACAGTACGCTGACGCTGGGACACCTCGGTGGTGGACATCATGG 184
 Db 347 CCGCGGCTACGATCATACAGGAGCGCTGGGACACCGGTCATCTCCGAGATCAAGG 406
 QY 185 -----GCGAGAGTCCGCTACTTCCACTCCATCCATCAAGAGGCGTGCACCGG 232
 Db 407 TCGCTGACGAGTACGAGAGGTGAGGTCTTCCACTGCTACAAAGCGGAGTGGACCGG 466
 QY 233 TGTGATTACACCCCTCGTTCCTGGCCAAAGTCTGGGGCAAGACCGGTCCTCAAGCTGT 292
 Db 467 TGTTCATACACCCCGTGGTTCCTGGAGAGGTCCCGGGCAAGATCCGAGGAGAGATCT 526
 QY 293 ACGGCCCGCGCTCGCGGCTGACTACTGAGAACACACAGCGCTTCGCCCTGTTCTGCA 352
 Db 527 ACGGCCCGACCGCGCACAGACTATGAGACAAACGAGCGCTTCAGCCTTCTCTGCC 586
 QY 353 AGGCGCGTATTGAGGCTGCCCGCTGCTGCCCTTC-----G 388
 Db 587 AGGAGCACTCGAGGCAACCGAGATCTCAACTCAACAAACACCCCTACTTTCTGGTC 646
 QY 389 GCCCGGCGAGGACTCGCTTCCTGGGCCAAGACTGGCACTCCGCCCTGGTGGCCGCTCC 448
 Db 647 CTAACGGGGAAGACGCTGGTGTGCTGCAACGACTGGCACACGCGGCTTCTGGCGCTGT 706
 QY 449 TGTGAGGACGAGTACACAGCCCAAGGGCGAGTTACCAAGGCCAAGTCTGGTGGCTA 508

Db 707 ACCTCAAGACCACTTACCAGTCCAAATGGCATCTACAGACGGCCAAAGTGGCTTCTGCA 766
 QY 509 TCACAACTATCGCTTCCAGGGCGGCATGTGGAGGAGGCTTTCAAGGACACAGAGTGC 568
 Db 767 TCCAAACATCTCGTACCAGGCGCTTCTCTTCGACGACTTTCGCGAGCTCAACCTGC 826
 QY 569 CCCAGGCGCGCTTTGACAAGTGGCTTCTCGGACGGCTATGCCAAGGTTTACACTGAGG 628
 Db 827 CCGACAGTTTCAAGTCGCTCTCGACTTCATTTGTCGCTACGACAAG-----873
 QY 629 CCACCCCATGAGGAGGACGAGAAAGCCCGCTGACGGGAAAGACCTACAAGAAGATCA 688
 Db 874 -----CCCGTGGAGGGCGC-----AAGATCA 895
 QY 689 ACTGGCTGAAGGTGGCATTATCGCCCGACAAAGCTGGTACTGTCTGCCCAACTACG 748
 Db 896 ACTGGATGAAGCGCGGATCTCGAGCGCCGACAAAGGTGCTGACGGTACGCCCTTACTACG 955
 QY 749 CGACCGAGATCGCTGCGCATGCCCGCGCGGTGTGGAGCTGGACACGCTCATCGCGCCA 808
 Db 956 CTGAGGAGCTCATCTCTGGCGAAGCCAGGGCTCGAGAGTCGACACATCATGCGCCTCA 1015
 QY 809 AGGCGATTGAGGCATTGTGAACGGCATGACATTTAGGAGTGGAAACCCCAAGACCGACA 868
 Db 1016 CGGGGATCACCGGAATCGTCAACGGCATGGACGCTAGTGGAGGAGCCCAAGGACA 1075
 QY 869 AGTTCTCTGTGCGCCCTACGACAGCAAGCGCTTACGCGCGCAAGCGCGCCCAAGG 928
 Db 1076 AGTTCTCGCGTCACTAGCATCATCACCGCGTGGAGGCGCAAGCGCTGAAACAGG 1135
 QY 929 AGGCCCTGACAGCGAGCTGGGCTGCTGTGACCCCAACCGCCCTGTTGCGCTTCA 988
 Db 1136 AGCGCTGACAGCGGAGTGGGCTGCGGTGGACAGAAAGTGGCGCTGGTGGCGCTTCA 1195
 QY 989 TCGCGCGCTGGAGGACAGAGGTGTGACATCATCTTGGCGCGCTGCCCAAGATCC 1048
 Db 1196 TCGGAGCTGGAGGACAGAGGCGCGGACGCTGATGATGCGCGCATCCCGAGATCC 1255
 QY 1049 TGGCCACCCCAAGTGCAGATGCCATCTCTGGTACCGCAAGCGCGCTACGAGAGC 1108
 Db 1256 TGAAGGAGGAGCTCCAGATCATCTCTTGGCACCGGAGAGAAAGTGTGAGAGC 1315
 QY 1109 TGTGTAACGCATCGGCACCAAGTACAAAGGCGCGCCCAAGGGCGTGGTCAAGTTCGG 1168
 Db 1316 TGTCTAAGACATGGAGGAGAAAGTTCGCGGCAAGGTGAGGGCGCTGCTCAGGTTCAAG 1375
 QY 1169 CGCCCTTGGCGACATGCTCACCGCGCGCGGCTTCACTGCTGGTGGCTCGCGCTTCG 1228
 Db 1376 CGCGCTTAGCTACAGATGATGGCGCGCGGCTTGTGCTGCTGACAGCGGCTTCG 1435
 QY 1229 AGCCCTGCGGCTGATTCAGCTGACGCCATGACGTACGCTACGCTGCGCGCTGATGCT 1288
 Db 1436 AGCCCTGCGGCTTATCCACGCTCCAGGATGCGCTATGGAACGCGCTGCTGCGCTT 1495
 QY 1289 CCACCGGCGCTTGTGTCGACACCGTCAAGAGGCGGCTCACCGGCTTCCATCGGGGCGCC 1348
 Db 1496 CCACCGGCGGCTGCTGTCGACACGATCGTGGAGGCAAGACCGGCTTCCATCGGGGCGCC 1555
 QY 1349 TGA-----ACCCCGACAACTGGACGAGCTGACGCCGAGCGCTTGGCGCGCACCGTGC 1402
 Db 1556 TCAGCTGCACTCAACGCTGGTGGAGCGCGGACCTGAAAGAGGTGGGACCACTCGTA 1615
 QY 1403 GCGCTGCCAGGAGTGTGGCGGCGGCTACCCGAGATGGTGGCACTGATCA 1462
 Db 1616 AGCGCGCTCAAGTGTGTCGCGACCGCGGCTACCGAGAGATGGTCAAGAACTCGGTGG 1675
 QY 1463 GCAGGACCTGCTCTGCTCCAAAGCCCGCCAGAAAGTGGAGGCGCTTCTGGAGAGGAGTGG 1522
 Db 1676 TCCAGGATCTCTCTGGAAGGACCTGCCAAGAACTGGGAGGACGCTGCTCTGGAAGTGG 1735
 QY 1523 TGTACGGAAGGCGCGCTGGGCCACCGCCCAAGAGAGAG 1560


```
QY 1523 TGTACGGCAAGCGCGGTGGCCACCGCCCAAGAGGAG 1560
Db 1751 GGGTGGAGGGAGCGAGCGGGGATCGTGGCGGAGGAG 1788

RESULT 15
AF486517
LOCUS AF486517 1827 bp mRNA linear PLN 02-JUL-2002
DEFINITION Hordeum vulgare cultivar Iyatoma Mochi granule bound starch
synthase I mRNA, complete cds.
ACCESSION AF486517
VERSION AF486517.1 GI:21667433
KEYWORDS
SOURCE
ORGANISM Hordeum vulgare.
Hordeum vulgare
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 1827)
AUTHORS Patron,N., Smith,A., Fahy,B., Hylton,C., Naldrett,M., Rossnagel,B.
and Denyer,K.
TITLE A mutation in the 5'-non-coding region of the barley GBSSI gene
alters its temporal and spatial expression and reduces GBSSI
activity and amylose content in the endosperm
JOURNAL Unpublished
AUTHORS Patron,N.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-2002) Metabolic Biology, John Innes Centre,
Norwich Science Park, Norwich, Norfolk NR4 7UH, UK
FEATURES
source
1..1827
/organism="Hordeum vulgare"
/cultivar="Iyatoma Mochi"
/db_xref="taxon:4513"
1..1827
/note="alpha glucosyl transferase"
/codon_start=1
/product="granule bound starch synthase I"
/protein_id="AA074050.1"
/db_xref="GI:21667434"
/translation="MAALATSQLATSGTLGVTDORSAPMFHAGFQGLRPKPADGT
FGMTVGASAPKSRKRAHNRKLSVVRATGSGMNLVFGAPMAPWSKTLGLGDV
LGGLPAPAAHGRVMVSPRYDQIKDAWDSVISEIKVADEYEVRFHCYKRGVDR
VFIDHPFLERKRGTKKEIKYDPAGTDYEDNQRFSLCQAALAPRILNINNPYF
SGPYGDFVFCNDWHTGLACYLKSNYSNGIYRTAKVAFCHINISYOGFSPDFA
QLNLPDRFKSDFIDGYDKPEGRKINWKAGILOADKVLTVSPVAEELISGEARG
CELDNIMRLTGITGVNMDVSEMDPTKDFLAVNYDITTLAKALNKEALOREVLG
PVDWKVPLVAFIGRLEQKGPDMVAVIPELLKEEDVQIILLGTGKKRFEKLLSMEE
KFPGKRVAVRFPNPLAHOMMAGADLLAVTSRFPCCGLIQGMRYGTPCVCASTGGL
VDTLVEGTFHFMGRSLSDCNVVEPADVKVATTLKRAVKVVGTPAYOEMYNKMIQD
LSWKGPAKNMEDVLLLELGVESGEPGIVGEETAPIAMENVAAP"
BASE COUNT 364 a 566 c 601 g 296 t
ORIGIN

Query Match 33.0%; Score 525.2; DB 8; Length 1827;
Best Local Similarity 61.5%; Pred. No. 3.9e-49;
Matches 982; Conservative 0; Mismatches 523; Indels 93; Gaps 5;

QY 5 TGGACATCGTGATGCTGCTGAGGTGCGCCCTGTGTCCAAAGACGGCGGCGTGGCG 64
Db 242 TGAACCTGCTGTCGCGCGCGAGATGGCGCCCTGGAGCAAGACCGCGGCGCTTGGCG 301
QY 65 ATGTGACTGTGGCTGCTATTGAGCTGGTCAAGCGGGCCACCGCGTCATGACCAATG 124
Db 302 ACCTGCTCGCGGCGCTTCCACCGCATGGCGGCCAACGGTCACGGGTGATGGTCTCT 361
QY 125 CCCCTCGTAGACCATGCTAGCGCTGGGACACCTCGGTGTCGTGGACATCATGG 184
Db 362 CCCCGCGCTACGATACATACAGGACGCGCTGGGACACCGAGGCTATCTCCGAGATCAAGG 421
QY 185 -----GCAGAAAGTCCGCTACTTCCACTCCATCAAGAAAGGGCGTGCACCGCG 232
```

```
Db 422 TCCTGACGAGTACGAGAGGGTGAAGTTCTTCCACTGCTACAAAGCGGAGTGACCGCG 481
QY 233 TGTGGATTGACCACCCCTGTTCTCTGTCGAAGGTCTGGGGCAAGACCGGCTCAAGCTGT 292
Db 482 TGTTCATCGACCACCCGCTGTTCTTGGAGAAAGTTCGGGGCAAGACCAAGGAAGATCT 541
QY 293 ACGGCCCCCGCTCCGGCGGTGACTTGTGACAAACACCAAGCGCTTCGCCCTGTCTTGCA 352
Db 542 ACGGGCCCGACGCGCGGACTACGAGACAAACAGACGCGCTTCAGCCTTCTCTGCC 601
QY 353 AGGCCGCTATGAGGTGCGCGGTGCTGCCCTTC-----G 388
Db 602 AGGCAGCGCTCGAGGACCCACGAGATCTCAACCTCAACACAAACCCCTACTTTCTGGTC 661
QY 389 GCCCGCGGAGGACTCGCTTCTGTCGCAACGACTGGGACTCGCCCTGGTCCCGTCC 448
Db 662 CCTACGGGAAGACGTGGTGTCTGTGCAACGACTGGACACGGGCTTCTGGCCTGT 721
QY 449 TGTGTAAGACGAGTACCAGCCCAAGGCCAGTTTCCAAAGGCCAAAGTCTGGTCTGCTA 508
Db 722 ACCTCAAGAGCAACTACCACTGTCATGATGCAATGAGGAGGCCAAGGTGGCTTCTGCA 781
QY 509 TCCACAACATCGCTTCCAGGGCCGATGTTGGAGGAGGCTTTCAAGGACACAAAGCTGC 568
Db 782 TCCACAACATCTGTTACAGGGCGCTTCTCCTTTCACGACTTTGCGCAGCTCAACCTGC 841
QY 569 CCGAGCGCGCTTTTGACAAAGTGGCTTCTCGGACGCTATGCCAAGGTTTACACTGAGG 628
Db 842 CCGACAGTTTCAAGTCTGCTTCTGACTTCTATGACCGCTACGACAG-----898
QY 629 CCACCCCATGGAGGAGGACGAGAACGCCCGCTGCGGSAAGACCTACAAGAAGATCA 688
Db 889 -----CCCGTGGAGGGGCGC-----AAGATCA 910
QY 689 ACTGGCTGAAGGTCGATTTATGCGCCCGCACAAGTGGTGTGACTGTGTGCCCAACTAG 748
Db 911 ACTGGATGAAGGCGGGATCTCTGAGGCGGACAAAGTGTGAGGTGAGCGCTACTACTAG 970
QY 749 CGACCCGAGATCGCTCGGATCGCGCGGGTGTGAGCTGGACACCGTCAATCCGCGCCA 808
Db 971 CTGAGGAGCTCATCTCTGGCAAGCCAGGGGTGCGAGTCTGACAAACATCATCGCGCTCA 1030
QY 809 AGGGCATTTGAGGCGCATTTGTAAGCGCATGACATTTGAGGAGTGGAAACCCCAAGACGACA 868
Db 1031 CGGGATCATCCGGAATCTGCAACGGCATGACGCTAGTGGTGGACCCCAAGACGACA 1090
QY 869 AGTTCTGTCTGGCGCTACGACCAAGAGAGGTCTACGCGGGCAAGCGCGCGCGCAAG 928
Db 1091 AGTTCTGTCTGCTCACTACGACATCACCGCGCTGGAGGCCAAGCGGCTGAACAAG 1150
QY 929 AGGCCCTGACGCGCGAGTGGGCTCCCTGTGACCCCGCGCGCGCTTTCGCTTCA 988
Db 1151 AGGCACTGACGCGCGAGTGGGCTGCGCGTGGAGTGGCGCTGGTGGCGCTTCA 1210
QY 989 TGGCGCGCTGAGGAGCAGAAAGGTTGTGGACATCATCTTGGCGCGCTTSCCAAGATCC 1048
Db 1211 TGGCGAGGCTGAGGAGCAGAAAGGCGCGCGTGTGATGATGCGCGTCAATCCCGGAGATCC 1270
QY 1049 TGGCCACCCCGCAAGTGCAGATCGCATCTCTGGTGTACGGCAAGCGCGCTACGAGAAGC 1108
Db 1271 TGAAGGAGGAGGACGTCACAGATCATCTCTTGGCACCGGGAAGAGTTTGAAGC 1330
QY 1109 TGGTGAACGCCATCGGCAAGTACAAGGCCGCGCAAGGGCGTGTCAAGTTCTCGG 1168
Db 1331 TGCTCAAGCAATGAGGAGAAAGTTCCCGGGCAAGGTGAGGGCCGTGGTTCAGTTCAAG 1390
QY 1169 GCGCCCTGGCGCACATGCTTCAACCGCGCGGCGGACTTTCATGCTGGTGGCGCTCGCGCTCG 1228
Db 1391 CGCGCGCTAGCTCACCAGATGATGGCGCGCGGAGCTTGTCTGCTGTCAACACCGCGCTCG 1450
QY 1229 AGCCCTGGGCGCTGATTCAGCTGACGCCATGCTACGTACCGGTACCGTGGCGGTAGCGT 1288
```

Db	1451	ACCCCTCGGGCGCTCATCAGCTCCAGGGATGCCTATGGAACCCGCTGCGGTGCGCGT	1510
QY	1289	CCACCGCGGCGCTGGTGACACCGTCGAAGGAGGGCGCTACCGCGTTCACATATGGGCGCC	1348
Db	1511	CCACGGGGGCTCTGCACACGATCTGGAGGCAAGACCTGTTTCCACATGGGCGCG	1570
QY	1349	TGA-----ACCCGACAAGCTGCAGGAGCTGACGCCGACGCCCTGGCGGCCACCGTGC	1402
Db	1571	TCAGCGTCGACTGCAACTGTTGGAGCGCGCGACGTGAAGAGGTGGCGACCAACCTGA	1630
QY	1403	GCCGTGCCAGCGAGTGTTCGGGGCGGCGCTACCCCGAGATGGTGGCCAAATGCATCA	1462
Db	1631	AGCGGCGGCTCAAGTCTCGCGCACGGCGGTACCAGAGATGGTCAGAACTGCATGA	1690
QY	1463	GCAGAGACTGTCCTGGTCCAAAGCCGCCCAAGTAGTGGAGGGCGCTCTGGAGAGGTGG	1522
Db	1691	TCCAGGATCTCTCTGGAAGGGACCTGCCAAGAACTGGGAGAGGTGCTTCGGAACTGG	1750
QY	1523	TGTACGGCAGGGCGCGTGGCCACCGCCAAAGAAGAG	1560
Db	1751	GGGTGGAGGGAGCGACCGCGGGATCTGTCGCGGAGAG	1788

Search completed: June 3, 2003, 13:34:51
Job time : 2792.82 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2003, 05:08:24 ; Search time 231.514 Seconds
(without alignments)
15495.523 Million cell updates/sec

Title: US-09-980-771A-8

Perfect score: 1593

Sequence: 1 gcgcgtgacatgctatggt.....ccgttgccgagaagatcccc 1593

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

```

1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	474.8	29.8	2267	18	AA63355
2	470	29.5	1915	19	AA63355
3	447.4	28.1	2542	19	AAV29753
4	344	21.6	1696	22	AA63355
5	316.8	19.9	2161	24	ABK53210
6	298	18.7	2115	24	AA63355
7	258	16.2	2183	24	ABK53215
8	221	13.9	2807	21	AA63355
9	221	13.9	2842	21	AA63355

10	218.2	13.7	2826	19	AAV01528	Wheat granule-boun
11	216.8	13.6	2097	19	AAV29755	Zea mays soluble s
12	214.8	13.5	2423	19	AAV70958	DNA encoding maize
13	214	13.4	2107	21	AA63355	Wheat starch synth
14	214	13.4	2107	21	AA63355	Wheat starch synth
15	212.4	13.3	9024	21	AA63355	Wheat SSII gene SE
16	210.2	13.2	2380	19	AAV66834	Zea mays soluble s
17	210.2	13.2	2920	24	ABK88112	CDNA encoding modi
18	209.2	13.1	2478	18	AA63355	Soluble starch syn
19	209.2	13.1	2939	21	AA63355	Wheat starch synth
20	209.2	13.1	2939	21	AA63355	Wheat starch synth
21	208.6	13.1	2946	24	ABK88115	CDNA encoding barl
22	208.6	13.1	2951	24	ABK88114	CDNA encoding barl
23	208.2	13.1	2950	24	ABK88113	CDNA encoding barl
24	207	13.0	2248	21	AA250651	Corn starch synth
25	206.6	13.0	1798	21	AA250647	Corn starch synth
26	206.6	13.0	2019	21	AA250646	Corn starch synth
27	203.8	12.8	1724	24	ABA01821	Rice starch synth
28	161	10.1	2007	19	AAV70959	DNA encoding maize
29	161	10.1	2007	19	AAV29754	Zea mays soluble s
30	161	10.1	2085	18	AA63355	Soluble starch syn
31	161	10.1	2085	19	AAV66833	Zea mays soluble s
32	151	9.5	1479	24	ABO90258	M. capsulatus gene
33	151	9.5	1479	24	ABO90406	M. capsulatus gene
34	125.6	7.9	333	21	AA63355	Human ORFX ORF124
35	125.6	7.9	333	21	AA63355	Human ORFX ORF124
36	118.2	7.4	2239	19	AAV01527	Wheat soluble star
37	118.2	7.4	2805	21	AA24487	Wheat soluble star
38	116	7.3	2662	20	AA63355	CDNA sequence of w
39	110.6	6.9	2533	15	AA45183	Soluble rice starc
40	108	6.8	1758	17	AA45183	Soluble starch syn
41	106.4	6.7	1984	21	AA63355	Arabidopsis thalia
42	102.6	6.4	1528	21	AA250637	Corn soluble starc
43	102.6	6.4	1620	19	AAV29759	Zea mays pEX52 st
44	102.6	6.4	1749	19	AAV70960	DNA encoding maize
45	102.6	6.4	1752	19	AAV29756	Zea mays soluble s

ALIGNMENTS

RESULT 1
AA63355
ID AAX63355 standard; cDNA; 2267 BP.
AC AAX63355;
XX AAX63355;
XX AAX63355;
DT 16-JUL-1999 (first entry)
XX Granule bound starch synthase encoding cDNA.
DE Maize; corn; Zea mays; delta-9 desaturase; GBSS; target; substrate;
KW granule bound starch synthase; hammerhead ribozyme; hairpin ribozyme;
KW modulation; gene expression; transgenic plant; cleavage; canola plant;
KW caffeine synthesis; coffee plant; nicotine production; tobacco;
KW fruit ripening; flower pigmentation; lignin production; ss.
OS Zea mays.
XX WO9710328-A2.
XX WO9710328-A2.
PD 20-MAR-1997.
XX 12-JUL-1996; 96WO-US11689.
XX 13-JUL-1995; 95US-0001135.
XX (DOWC) DOWELANCO.
PA (RIBO-) RIBOZYME PHARM INC.
XX Edington BE, Folkerts O, Guo L, McSwiggen JA, Merlo DJ;
PI Merlo PAO, Skokut TA, Young SA, Zwick MG;


```

1619 TCCAGCGGCCATCAAGTGTTCGCGACGCGCGCTACGAGGAGATGTTGAGAACTGCA 1678
1460 TCAGCCAGGACTGCTCTGGTCCRAAGCCGCGCCAGAGTGGAGGCGCTGCTGGAGGAG 1519
1679 TGATCCAGGATCTCTCTGGAGGCGCTGCCAAGAACTGGAGAACGCTGCTGCTAGCC 1738
1520 TGGTGTAGCGCAAGGCGCGCTGGCCACCGCCCAAGAGGAGGATCAAGGTGCGCGCTTG 1579
1739 TCGGGGTGCCC---GGCGGCGAGCCAGGGTTCGAGGCGGAGAGATCGCGCGCTCGCA 1795
1580 CCGAAGATCCCC 1593
1796 AGGAGAAGCTGGCC 1809

```

RESULT 3

AAV29753
ID AAV29753 standard; DNA; 2542 BP.

AC AAV29753;

DT 11-SEP-1998 (first entry)

XX Oryza sativa waxy gene.

DE SER; starch-encapsulating region; fusion vector;

KW starch synthase; bacterial glycoen; ss.

XX Oryza sativa.

Key Location/Qualifiers

PH 453..2282

FT CDS

FT /*tag= a

FT /product= starch (bacterial glycoen) synthase

FT polyA_site

FT 2535

FT /*tag= b

XX WO9814601-Al.

XX 09-APR-1998.

XX 30-SEP-1997; 97WO-US17555.

XX 30-SEP-1996; 96US-0026855.

XX (EXSE-) EXSEED GENETICS LLC.

XX Guan H, Keeling P;

XX WPI; 1998-240100/21.

XX P-PSDB; AAW56485.

XX Hybrid polypeptide comprising starch-encapsulating region and protein - useful for, e.g. producing protein(s) resistant to degradation by stomach acids

XX Example 2; Page 32-34; 156pp; English.

XX The sequence is that of the waxy gene which codes for starch synthase. It can be used in the production of a hybrid polypeptide comprising a starch-encapsulating region (SER) fused to a payload protein. The hybrid polypeptide can be used to make modified starches comprising the payload protein, selected from, e.g. hormones, growth factors, antibodies, enzymes, dyes, immunoglobulins, etc. The modified starch can also be used to provide grain feeds enriched in amino acids. By encapsulating the payload protein in starch, it is more resistant to degradation by stomach acids.

XX Sequence 2542 BP; 610 A; 665 C; 693 G; 574 T; 0 other;

XX Query Match 28.1%; Score 447.4; DB 19; Length 2542;

Best Local Similarity 58.6%; Pred. No. 1.2e-61; Matches 930; Conservative 0; Mismatches 561; Indels 96; Gaps 5;

```

QY 5 TGGACATCTGATGTTGCTGAGTCCGCCCTTGGTCCAAGACGGCGGCTGGGG 64
DB 700 TGAACGTCGTGTTGTCGGCGCGGAGATGCCCTTGGAGCAGACCGCGGCTGG 759
QY 65 ATGTGACTGTGGCTGCTGCTATTTAGCTGGTCAAGCGCGGCCACCGCTCATGACCAT 124
DB 760 AGTCTCTGCTGGCTCCCGCTCCCGCTGCCATGGTGCGAATGCCACAGGGTCTGGT 819
QY 125 CCGCTCGTACGACAGTACGCTGACGCTGGGACACCTCGGTGGTCTGGACATCATGG 184
DB 820 CTCCTCGTACGACAGTACGACGCTGGGATACCAAGCGTGTGGGTGAGATCAAGG 879
QY 185 -----CGAGAAGGTCGCTACTTCCACTCCATCAAGAAGGCGGTGACCGCG 232
DB 880 TTGCACAGAGTACGAGAGGTTGAGGTTTTTCATTTGCTACAAGCTGGAGTGCACG 939
QY 233 TGTGATTGACCAACCGCTGTTCTTGGCCAAAGTCTGGGSCAAAGACCGGTCTCAAGTGT 292
DB 940 TGTTCATCGACCATCGTCTATCTCTGGAAGGTTTGGGAAAGACCGGTGAGAAGATCT 999
QY 293 ACGGCGCGCTCCGCGGCTGACTACTGACAAACACAAAGCGCTTGGCCCTGTTCTGCA 352
DB 1000 ACGGACCTGACACTGGAGTTGATTACAAGACAAACAGATGCGTTTACGCTTCTTTGCC 1059
QY 353 ACGCGCTATTGAGGCTGCGCGGCTGCTGCCCTTCGCGCC----- 393
DB 1060 AGCAGCACTCGAGGCTCTAGATCTTAACCTTAAACACACCACTACTTCAAAGGAA 1119
QY 394 -----GGCAGGACTGCTGCTTCTGTTGGCCAAAGCTGGGACTCCGCCCTGTTGCCGTCC 448
DB 1120 CTTATGTTGAGGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1179
QY 449 TGTGAAGACAGTACGACGCGCGGCGGCGGCTTCAAGGACCAAGTCTGCTGCTGCTGCT 508
DB 1180 ACCTGAAGAACAACTACCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1239
QY 509 TCCACAACATCGCTTCCAGGCGCGGCTGTTGGAGGAGGCTTCAAGGACCAAGTCTGCTGCTGCT 568
DB 1240 TCCACAACATCTCTACCGGCGCGGCTTCTGCTTTCGAGGATTACCTGAGCTGAACCTCT 1299
QY 569 CCGAGCGCGCTTGTACAGCTGCTGCTTCTGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 628
DB 1300 CCGAGAGGTTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1343
QY 629 CCACCCCGCTGAGGAGGACGAGAGCGCGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 688
DB 1344 -----ACGCCGTTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1368
QY 689 ACTGGCTGAAGGTTGCTATTCGCGCGCGGACAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 748
DB 1369 ACTGGATGAAGCGCGGAACTCTGGAAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1428
QY 749 CGACCCAGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 808
DB 1429 CCGAGGAGCTATCTCTCGGATCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1488
QY 809 AGGCGATTGAGGCGGCTTGTGAAGGCTGACATTTAGGAGTGGAAACCCCAAGACCGGAC 868
DB 1489 CCGGATCACCGGCTGCTCAACGCGGATGAGCTGACGAGTGGGATCTCTAGCAAGGAGCA 1548
QY 869 AGTTCCTGCTGCGCGCTTACGACCAAGACAGCTCTACGCGCGGAGGAGGAGGAGGAGGAGGAGGAG 928
DB 1549 AGTACATACCGCGGAACTACGACCAACCGGCAATTCGAGGCGGAGGAGGAGGAGGAGGAGGAGG 1608
QY 929 AGGCGCTGAGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 988
DB 1609 AGGCGTTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1668
QY 989 TCGGCGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1048

```

Db 1669 TCGCGAGGCTGGAGAACAGAGGGCCCTGACGTGTCATGCGCGCCGACCTCCGGAGCTCA 1728
 QY 1049 TGGCCACCCCAAGGTGTCAGATGCCATCTGGGTACCGCAAGCGCGCTACGAGAAGC 1108
 Db 1729 TG---CAGGAGGACGTCACATGCTTCTTGGGTACTGGAAGAAAGAACTTCGAGAAGC 1785
 QY 1109 TGGTGAAGCCATCGCGACCAAGTACAAAGGCGCGCCCAAGGGCGTGGTCAAGTTCCTCG 1168
 Db 1786 TGTCTAAGACATCTGAGGAGAGATATCCGGGCAAGGTGAGGGGGTGGTGAAGTTCAACG 1845
 QY 1169 CGCCCTTGGCGCATGCTCACCGCGCGCGCCGACCTTCATGCTGGTGCCTCGCGCTTCG 1228
 Db 1846 CGCGCTTGTCTCATCTCATGTCGCGGAGCGACGCTGCTCGCGTCCCGACGCGCTTCG 1905
 QY 1229 AGCCCTCGGCGCTGATCCAGCTGACGCCATGACCTACGCTACCGTACCGTCCCGTGGTACGCT 1288
 Db 1906 AGCCCTGTGGACTCATCCAGCTGACGGGATGAGATACGGAAGCGCCCTGTGCTTGCCTG 1965
 QY 1289 CCACCGCGCGCTTGTGTCACACCTGTCAAAGGAGGCGCTCACCGGCTTCACATGGCGCCC 1348
 Db 1966 CCACCGGTGGGCTGCTGGACACCGTCTATCAAGGCAAGACTGGTTTCCACATGGCGGCTC 2025
 QY 1349 TGAACCCCGAC-----AAGCTGAGCAGGCTGACGCGCGACGCGCTGGCGCGCCACCGTGC 1402
 Db 2026 TCAGCGTCGACTCAAGGTGGTGGAGCCAAAGCGACGCTGAAGAGGTGGCGGCCACCTGA 2085
 QY 1403 GCGTGCCAGCAGAGTGTGTTGCGGCGGCGCGCTACCCCGAGATGTTGGCGCACTGCATCA 1462
 Db 2086 AGCGGCCCATCAAGGTGCTGCGGACGCGCGCTACGAGGAGATGGTCAGGAAGTGCATGA 2145
 QY 1463 GCAGGACCTGTCTGTGTCCTCAAGCGCGCCCGCAGAAAGTGGGAGGCGCTGCTGGAGGAGTGG 1522
 Db 2146 ACCAGGACCTCTCTTGGAGGGGCGCTCGGAGAACTGGGAGATGCTCTCTGGCGCTGG 2205
 QY 1523 TGTACGGCAAGGCGCGGTGGCCACCG 1549
 Db 2206 GCGTCGCGGACGCGCGGGGATCG 2232

RESULT 4
 AAC86955
 ID AAC86955 standard; cDNA; 1696 BP.
 AC AAC86955;
 DT 02-APR-2001 (first entry)
 XX cDNA encoding a granule bound starch synthetase II (GBSSII).
 DE Granule bound starch synthetase II; GBSSII; starch grain;
 KW adenosine diphosphate glucose-alpha,4-glucan alpha4-glucosyltransferase;
 KW starch synthetase; ds.
 OS Chlamydomonas reinhardtii.
 FH Key Location/Qualifiers
 FT CDS 3..719
 FT /*tag- a
 FT /*product= "granule bound starch synthetase II (GBSSII)"
 XX
 XX FR2793806-A1.
 XX
 XX 24-NOV-2000.
 XX
 XX 21-MAY-1999; 99PR-0006494.
 XX
 XX 21-MAY-1999; 99PR-0006494.
 XX
 XX (CNRS) CNRS CENT NAT RECH SCI.
 XX
 XX Dhulst C, Ball S;
 XX
 XX WPI; 2001-052291/07.
 DR

DR P-PSDB; AAB311175.
 XX New recombinant nucleic acid encoding fusion of starch synthase and
 PT second component, useful in pharmaceutical and food compositions, is
 PT targeted to starch granules.
 XX Claim 3; Fig 1; 52pp; French.
 PS
 XX
 CC The present sequence encodes a granule bound starch synthetase II
 CC (GBSSII). The specification describes a recombinant nucleic acid,
 CC comprising a polynucleotide encoding an adenosine diphosphate
 CC glucose-alpha,4-glucan alpha4-glucosyltransferase or starch
 CC synthetase, placed upstream of a sequence that encodes a polypeptide
 CC of interest. The adenosine diphosphate glucose-alpha,4-glucan
 CC alpha4-glucosyltransferase protein can migrate to sites of
 CC biosynthesis of starch grains in plant cells, becoming associated
 CC with these grains. The recombinant nucleic acid sequence is used to
 CC target polypeptides of interest to starch grains.
 XX
 SQ Sequence 1696 BP; 347 A; 509 C; 571 G; 269 T; 0 other;
 Query Match 21.6%; Score 344; DB 22; Length 1696;
 Best Local Similarity 100.0%; Pred. No. 1.8e-45;
 Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1250 TGCACGCCATGCTACGTACCGTACCGTGGCGGCTGAGCTCCACCGCGGCTGTGTCGACA 1309
 Db 13 TGCACGCCATGCTACGTACCGTACCGTGGCGGCTGAGCTCCACCGCGGCTGTGTCGACA 72
 QY 1310 CCGTCAGGAGGGCGTCCACGGCTTCACATGGCGGCCCTGAACCCCGACAAGTGGACG 1369
 Db 73 CCGTCAGGAGGGCGTCCACGGCTTCACATGGCGGCCCTGAACCCCGACAAGTGGACG 132
 QY 1370 AGGCTGACGCGCGACGCGCTTGGCGGCCACCGTGGCGGCTGAGCTGTTCGCGGCG 1429
 Db 133 AGGCTGACGCGCGACGCGCTTGGCGGCCACCGTGGCGGCTGAGCTGTTCGCGGCG 192
 QY 1430 GCGCTACCGCGGAGATGGTGGCCAACTGCATCAGCCAGGACCTGTCTGTGTCGAAGCCCG 1489
 Db 193 GCGCTACCGCGGAGATGGTGGCCAACTGCATCAGCCAGGACCTGTCTGTGTCGAAGCCCG 252
 QY 1490 CCCAGAAAGTGGAGGCGCTGCTGAGGAGTGGTGTACGGCAAGGGCGGCTGGCCACCG 1549
 Db 253 CCCAGAAAGTGGAGGCGCTGCTGAGGAGTGGTGTACGGCAAGGGCGGCTGGCCACCG 312
 QY 1550 CCAGAAAGGAGGAGATCAAGTGGCGGCCCTTGGCGAAGAGATCCCC 1593
 Db 313 CCAGAAAGGAGGAGATCAAGTGGCGGCCCTTGGCGAAGAGATCCCC 356

RESULT 5
 ABK53210
 ID ABK53210 standard; DNA; 2161 BP.
 XX
 AC ABK53210;
 XX
 DT 14-AUG-2002 (first entry)
 XX
 DE Potato granule-bound starch synthase I (GBSSI) DNA sequence.
 XX
 XX Starch; amylose; cereal crop; commodity starch; thickened foodstuff;
 KW bread; baking; pasta; potato; granule-bound starch synthase I; GBSSI; ds.
 XX
 OS Solanum tuberosum.
 XX
 XX WO200218606-A1.
 XX
 XX 07-MAR-2002.
 XX
 XX 28-AUG-2000; 2000WO-US23494.
 XX
 XX 28-AUG-2000; 2000WO-US23494.
 XX
 XX

PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Broglie K, Lightner J;
XX
DR WPI; 2002-404556/43.
XX
PT Novel starch isolated from cereal crop with a modification of amylose
PT content and/or amylose weight distribution relative to commodity starch
PT from same species of the cereal crop, for preparing thickened
PT foodstuffs
XX
PS Example 1; Page 36; 44bp; English.
XX
CC The present invention relates to a new starch comprising amylose,
CC where the starch is isolated from a cereal crop, and comprises a 50%
CC increase in average molecular weight relative to commodity starch from
CC the same species of the cereal crop from which starch was isolated.
CC The starch of the invention is useful for preparing thickened foodstuffs
CC by combining the starch, a foodstuff and water and cooking the resulting
CC composition as necessary to produce a thickened foodstuff. The invention
CC is also useful for preparing foodstuffs such as breads, baked goods,
CC pastas, etc. The present nucleic acid sequence represents the potato
CC granule-bound starch synthase I (GBSSI) DNA sequence that was used in
CC the methods of the invention for generation of a potato full-length
CC sense GBSSI DNA construct.
XX
SQ Sequence 2161 BP; 614 A; 428 C; 505 G; 614 T; 0 other;

Query Match 19.9%; Score 316.8; DB 24; Length 2161;
Best local Similarity 53.8%; Pred. No. 3.2e-41;
Matches 836; Conservative 0; Mismatches 622; Indels 96; Gaps 5;

QY 5 TGGACATCGTATGTTGCTGCTGAGTCCGCCCTGTGCTCCAGACGGCGGCTGGCG 64
DB 278 TGAACCTGATCTTTGGGGTACTGAGTGTGCTTGGACAAACAGTGGTACTAGTG 337
QY 65 ATGTGACTGGTGGCTGCTATGAGTGGTCAAGCGCGGCCACCGCTCATGACCAATG 124
DB 338 ATGTTCTTGGTGGATACCAACAGATGCTTGGGATCTAGCGTTGGCGTTGAGGTCAAAG 397
QY 125 CCCTGCTTACACACGACGTGAGCGCTGGGACACCTCGTGGTCTGGGACATCATG 184
DB 398 CCCCCGTTATGACCAATCAAGATGCTTGGGATCTAGCGTTGGCGTTGAGGTCAAAG 457
QY 185 -----GCGAAGAGCCGCTACTTCCACTCCATCAAGAGGCGCTGACCGG 232
DB 458 TTGGAGACAGCATGAAATGTTGCTTCTTCACTGCTATAAAGCGTGGGTGATCGTG 517
QY 233 TGTGATTCACACCGCTGCTTCCGGCAAGGTCTGGGGCAAGACCGGCTCAAGCTGT 292
DB 518 TTTTGTGGACACCAATGTTCTTGGAGAAAGTTTGGGGCAAACTGGTTCAAAATCT 577
QY 293 ACGGCGCGCTCCGGCGCTGACTACTCGGACCAACCAAGCGCTTCGCCCTCTCTGCA 352
DB 578 ATGGCCCCAAGCTGGACTAGATTATCTGGACAAATGAACTAGGTTCAGTTGTTGTC 637
QY 353 AGCGCGCTATTGAGCTGCGCGCTGCTGCTGCTCGGC----- 390
DB 638 AAGCAGCCCTAGAGGCACTAAAGTTTGAATTTGAACAGTAGCAACTACTTCTCAGGAC 697
QY 391 ---CCGGCGAGGACGCTGCTTCTGGGCCAAGCACTGGCACTCCGCGCTGGTCCGCTCC 448
DB 698 CATATGGAGGAGGTGTTCTCTTCAATGCGCAATGATTGGCACAGCTCATCTTCTGT 757
QY 449 TGTGAAGGAGGAGTACCAAGCCCAAGGCGGCTTACCAAGGCGCAAGTCTGCTGGCTA 508
DB 758 ACTTGAAGTCAATGTACCAAGTCCAGAGAACTATTTGAATGCCAAGTCTGCTTCTGCA 817
QY 509 TCCACAAATTCGCTTCCAGGGCGCAATGTTGGAGGAGGCTTTCAAGGACACGAACTGC 568
DB 818 TCCATAACATTTGCTTACCAAGGCGGATTTTCTTCTCTGACTTCCCTCTTCTCAATCTTC 877
QY 569 CCAGGCGCGCTTTGACAACTGCGCTTCTCGGACGCGCTATGCCAAGGTTTACACTGAG 628

Db 878 CTGATGAATTCAGGGGTTCTTTTGATTTCATTCGATGATGAGAAGCCCTGTAAAG--- 934
QY 629 CCACCCCCATGGAGGAGGAGGAGAGCCGCCCTGACGGGAAAGACCTACAGAAGATCA 688
Db 935 -----GTAGGAAATCA 946
QY 689 ACTGGCTGAAGGTGGCATATTATCGCGCCGACAAAGCTGGTACTGTGTCGCCCAATACG 748
Db 947 ACTGGATGAAGCTGGGATATTAGAAATCATAAGGTGGTTACAGTGAGCCCATATATG 1006
QY 749 CGACCGGATCGCTGCCGATGCGCGCGGTGTGAGCTGGAGCTGGACACCGTCAATCCGCCCA 808
Db 1007 CCCAAGAACTGTCTCTGCTGTGTGACAAGGGTGTGAATTTGGACAGTGTCTTCTGTAAGA 1066
QY 809 AGGGCATTTAGGCGCATTTGAAGCGCATGGACATTCAGGAGTGAACCCCAACACCGACA 868
Db 1067 CTTGCACTACTGGGATTTGAATGGCATGATACACAGAGTGGAAACCCAGCGCTGACA 1126
QY 869 AGTCTCTGTCTGCGCCCTACGACCAAGACAGCTCTACGCCGCAAGCGCGCCCAAGG 928
Db 1127 AATACACAGATGTCAAATACGATATACCACTGTCTATGGACGCAAAACCTTTACTAAGG 1186
QY 929 AGSCCTTGAGCGCGAGCTGGGCTGCTGTGACCCCGCCCGCCCTGTTCCGCTTCA 988
Db 1187 AGGCTCTTCAAGCAGCAGTTGGCTTGGCTGTGACAAAGATCTCCCTTTGATTGGCTTCA 1246
QY 989 TCGGCGCGCTGGAGAGCAGAGGGTGTGGACATCATCTTGGCGCGCTCCCAAGATCC 1048
Db 1247 TCGGACAGCTTGGAGAGCAGAGAGTTTCAGATATTTCTTGTGCAATTCACAAGTTCA 1306
QY 1049 TGCCCAACCCCAAGTGCAGATCCCATCTCTGGTACCGGCAAGCGCGCTACGAGAAGC 1108
Db 1307 T---CGGATTTGATTTCAAATTTGTTAGTCTCTTGAATGSCAAAAGAGAGTTTGCAGG 1363
QY 1109 TGTGAACGCGCATCGGACCAAGTACAGGCGCGCCGCAAGGCGGTGCTCAAGTTCTCGG 1168
Db 1364 AGATTGAACAGCTCGAAGTGTGTACCCCTAACAAAGCTAAGAGGTGGCAAAATTCATG 1423
QY 1169 CGCCCTGGCGGCACTGTCCACCGCGCGCGCTTCACTGCTGGTGGCGCTCGCGCTCG 1228
Db 1424 TCCCTTTGGCTCAGATGATCAGTCTGCTGCTCATTTATGTTGGTTCCAAAGCAGATTG 1483
QY 1229 AGCCCTGCGGCTGTATCCAGCTGCAGCCCATGCACTACGTTACGCTGCGCGCTGGTAGCCT 1288
Db 1484 AACCTTGTGCTCAATTCAGTTACATGCTATGCGATATGGAACAGTGCCTAATCTGTGCAT 1543
QY 1289 CCACGCGCGCTGTGTCGACACCGTCAAGGAGGCGCTCACCGCTTCCACATGGCGGCC 1348
Db 1544 CGACTGTTGACTGTGTGACACTGTGAAGAGGCTATATCTGATTCATATGGAGGCTT 1603
QY 1349 TGA-----ACCCCGACAAAGCTGGAGGCTGACGCGCGAGCGCTGGCGGCCACCGCTG 1402
Db 1604 TCAATGTTGAATCGATGTTGTTGACCCAGCTGATGTGCTTAAGATAGTAAACAAGTTG 1663
QY 1403 GCGTGGCAGGAGTGTGTTGGGGGGCGGCTACCCCGAGATGGTGGCCCACTGCATCA 1462
Db 1664 CTAGAGCTCTTGGAGTCTATGGCACCCCTCGCATTTGCTGAGATGATAAAAAATTTGCATGT 1723
QY 1463 GCCAGGACCTGCTCTGTTCCCAAGCCCGCCAGAGTGGGAGGCGCTGCTGCGAG 1516
Db 1724 CAGAGGAACCTCTCTGGAAGGAACCTGCCAAGAAATGGGAGACATTTGCTATTGG 1777

RESULT 6
AAC32824
ID AAC32824 standard; DNA; 2115 BP.
XX
AC AAC32824;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 779.


```
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 18.7%; Score 298; DB 21; Length 2115;
Best Local Similarity 53.7%; Pred. No. 2.8e-38;
Matches 820; Conservative 0; Mismatches 610; Indels 96; Gaps 6;

QY 20 TTGCTGCTGAGTGCCTTGGTCCAGACGGGGCGCTGGCGGAGTGTGACTGTGGGCC 79
DB 415 TTGGAGCTGGAAGTTGGTCCATGAGTAACACUGGTGGTCTTGGTGAATTCGGGTTG 474
QY 80 TGCATTATTGAGTGTGAAGCGCGCCGCGTGCATGACCATGTCCTGCTGCTAGCACC 139
DB 475 TACCTCCAGCTCTTGTGCTAGAGGCCACCGTGTGATGACAAATTTGTCTCTGCTAGCC 534
QY 140 AGTAGCTGACGCTTGGACACCTCGGTGGTGTGCTGGACATC-----ATGGGCG 187
DB 535 AATATAAAGATGCTTGGGACCATTTGTGTGGTTTCAGATCAAAAGTTGGGGATTAAGTTG 594
QY 188 AGAAGTCCGCTACTTCCACTCCATCAAGAAGGGCGTGCACGCGTGTGGATTGACACC 247
DB 595 AGAATGTTCTTCTTCCATTCCTACAAACGAGGAGTTGATCGTCTTTGTGTGACCATC 654
QY 248 CTTGTTCTTCTGCGCAAGGTTCTGGGCAAGACCGGCTCCAAAGTGTACGCGCCCGCTCG 307
DB 655 CAATCTTTCTTGTAAAGTTGTGGCAAAACAGGATCCAAAATCTATGTCTCTATAACTG 714
QY 308 GCGCTGACTTACCTGGACACACAGCGCTTCGCCCTGTCTCTGCRAAGCGCGCTATTGAGG 367
```

```
DB 715 GAGTAGACTACAAATGACAAACCAACTCCGGTTCAGTTTGTGTCAAGGCTGCTCTTGAGG 774
QY 368 CTGCGCGGTGCTGCCCTTCGGCCCGGGGAGGACTGC-----405
DB 775 CACCACAGGTTCTGAACCTGAACAGCAGCAAGTACTCTCTCGACCAATATGTTGAAGATG 834
QY 406 --GTCTTCGTGGCCCAACGACTGGCCTCGCCTGTGTGCCCTCTCTGTGTGAAGGAGAGT 463
DB 835 TAGTCTTTTGGCAATGACTGSCACACTGCTCTACTTCCATGTTTACCTCAAACTCTATCT 894
QY 464 ACCAGCCCAAGGGCCAGTTCCACCAAGGCCAAGTCGGTGTGCTGCTATCCACACATCGCCT 523
DB 895 ATCAATCCCGCGGAGTCTACATGAATGCAAGGTGGTCTTCTGCAATCCAAACATAGCCT 954
QY 524 TCCAGGGCGCATGTGGGAGGAGGCTTTCAAGGACACGAAGCTTCCAGGACAGGAGGCGCCTTG 583
DB 955 ACCAGGAAGAT-----TGCTTTG 975
QY 584 ACAAGTGGCCTTCTGGACGGCTATGCCAAGGTTTACACTGAGGCCACCCCATGGAGG 643
DB 976 ATGACTATTCCCTTCTCAACTTGCOCATCAGCTTTTAAAGTTCTTTTCGACTTCATGGAG 1035
QY 644 AGGACGAGAACGCCCGCTGACGGGAAAGACCTACAGAGATCAACTGGCTGAAGGGTG 703
DB 1036 GGTATGAAAAGCC-----AGTAAAGGACGGAATAAATGATGGATGAAGGCTG 1083
QY 704 GCATTATCGCCGCGCAAGCTGTGTGCTGCTGCGCAACTACGCGACCGAGATCGCTG 763
DB 1084 CAATCTGGAAGCTACCGTGTCTTAACAGTTAGTCCATCTACTAGCTCAAGACTCATCT 1143
QY 764 CCGATGCCCGCGGTGTGGAGCTGGACACCGTCTATCCGCGCCCAAGGGCATTTGAGGGA 823
DB 1144 CTGGAGTTGATAGAGCGGTGGAATTCATAAATATTCGAATGAAACAGATTTCCGGAA 1203
QY 824 TTGTGAACGCGATGGACATTGAGAGTGAACCCCAAGACCGACAGTTCCTCTCTGCGC 883
DB 1204 TTATTAAATGGAATGGAATTCAGAATGAACCCCGTCTACTGCAAGTACATCGATATCA 1263
QY 884 CCTACACCAAGACAGCGTCTAGCGCGGAAAGCGCCGCCAGAGGAGGCGCTCAGGCGCG 943
DB 1264 AATACGATATTACCAGTGTACAGATGCTAAACCAATGATCAAGAAGACACTTCAGGCTG 1323
QY 944 AGCTGGCGCTGCTGTGGACCCCGCCCGCTGTTCGCTTCATCTCGCCGCTGAGG 1003
DB 1324 CTGTTGGACTTCCCGTGGACAGGATGTCCCGGTTATTCGGTTTCATAGGAGATTTGAGG 1383
QY 1004 AGCAGAGGGTGTGGACATCATCTCGCGCGCTGCGCAAGATCTCTGGCCACCCCAAGG 1063
DB 1384 AGCAGAGGGTGTGATATTCTAGTGAAGCTATTTCGAAGTTTCATGG---GGCTCAATG 1440
QY 1064 TGCAGATCGCCATCTCTGGTACCGCAAGSCCGCTACGAGAAGTGTGTAAGCGCATCG 1123
DB 1441 TTCAGATGGTTATCTCTGGACTTGAAGAAGAAGATGAGGCTCAGATTTCTTGAAGTAG 1500
QY 1124 GCACCAAGTACAAAGGCGCGCGCGCAAGGGGTGTCAAGTCTCGGCGCCCTCGGCGACA 1183
DB 1501 AGAGAAGTTTCCCAAGGAGGCGGTTCGAGTGGCGGAATTCACAGTGCATGCGCTCAT 1560
QY 1184 TGCACCGCGCGCGGCTTCATGCTGTGCTGCTCGCGCTTCGAGCGCTCGGCGCTGA 1243
DB 1561 TGATCACTGCTGGAGCTGACTTCATCTATGTCCTCAAGAGGTTTGAGCGCTGTGGTCTCA 1620
QY 1244 TCCAGTGCACGCGCATCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1303
DB 1621 TTCAGTGCACGCAATGATATGGAACCGTCCCTATTTGCGCATCTACTGTGGACTTG 1680
QY 1304 TGCACACCGTCAAGAGGGGTTCACCGGCTTCCACATGGG-----CGCCCTCAACCCCG 1357
DB 1681 TGGACACTGTGAAGATGCTACACAGGTTTCCACAGTTTGAAGATTCAGGTTCAAGTGTG 1740
QY 1358 ACAAGCTGACAGGCTGAGCGCGCGCTGCGCGCCACCGTGGCGCGCTGCGCAGCGAGG 1417
```

Db 1741 AAGTTGTGGATCCAGATGATGTGATGACACAGCAAGAGCTGTGACAGAGCCGTTCGAG 1800
 QY 1418 TCTTTTGGGGGGGGCGCTACCCCGAGATGGTGGCCAACTGCATCAGCCAGGACCTGTCT 1477
 Db 1801 TATATGGAAACATCCCAATGCAAGAAATGGTCAAGAACTGCATGACCAAGACTTCTCT 1860
 QY 1478 GTCCCAAGCCCGCCAGAGTGGAG 1503
 Db 1861 GGAAGGACCTGCGAGGTGTGGAG 1886

RESULT 7
 ABK53215
 ID ABK53215 standard; cDNA; 2183 BP.
 AC ABK53215;
 XX
 XX
 DT 14-AUG-2002 (first entry)
 XX
 DE Canna edulis granule-bound starch synthase I (GBSSI) cDNA sequence.
 XX
 KW Starch; amylose; cereal crop; commodity starch; thickened foodstuff;
 KW bread; baking; pasta; potato; granule-bound starch synthase I; GBSSI; ss.
 OS Canna edulis.
 XX
 XX WO200218606-A1.
 PN 07-MAR-2002.
 XX
 XX 28-AUG-2000; 2000WO-US23494.
 XX
 XX 28-AUG-2000; 2000WO-US23494.
 PR
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 FA
 XX Broglie K, Lightner J;
 PI WPI; 2002-404556/43.
 XX
 DR
 XX Novel starch isolated from cereal crop with a modification of amylose
 PT content and/or amylose weight distribution relative to commodity starch
 PT from same species of the cereal crop, for preparing thickened
 PT foodstuffs -
 XX
 PS Example 1; Page 37-38; 44pp; English.
 XX
 CC The present invention relates to a new starch comprising amylose,
 CC where the starch is isolated from a cereal crop, and comprises a 50%
 CC increase in average molecular weight relative to commodity starch from
 CC the same species of the cereal crop from which starch was isolated.
 CC The starch of the invention is useful for preparing thickened foodstuffs
 CC by combining the starch, a foodstuff and water and cooking the resulting
 CC composition as necessary to produce a thickened foodstuff. The invention
 CC is also useful for preparing foodstuffs such as breads, baked goods,
 CC pastas, etc. The present nucleic acid sequence represents the Canna
 CC edulis granule-bound starch synthase I (GBSSI) cDNA sequence in clone
 CC ectlc.pk007.015 that was used in the methods of the invention for
 CC generation of a full-length sense Canna edulis GBSSI construct.
 XX
 SQ Sequence 2183 BP; 636 A; 428 C; 516 G; 603 T; 0 other;
 XX

Query Match 16.28; Score 258; DB 24; Length 2183;
 Best Local Similarity 51.68; Pred. No. 5.1e-32;
 Matches 794; Conservative 0; Mismatches 650; Indels 96; Gaps 5;
 QY 5 TGGACATCGTATGTTGCTGCTGAGGTGCGCCCTTGGTCCAAAGACGGCGCTGGGCG 64
 Db 357 TGAACCTGGGTTTGTGGTGTGAGGTAGCTCCATGAGCAAACTGGGGCCCTTGGCG 416
 QY 65 ATGTGACTGGTGGCTGCTATGAGCTGGTCAAGCGCGGCCACCGCTCATGACATTG 124
 Db 417 ATGTTCTTATAGGATTGCCACCTGCTATGCTGCTGCAATTGGGCACAGGCTCATGCCGTGG 476

QY 125 CCCCTGCTACGACCACTAGCTGAGCCCTGGGACACCTCGGTGTCGTGGACAT----- 179
 Db 477 CGCCAGATATGACCAATATAAAGATATCTGGATACAAAGTGTCCAGTTGAGTTAAAG 536
 QY 180 -----CATGGCGGAGAAGGTCCGCTACTTCGACTCCATCAGAAAGGGGTGCACCGCG 232
 Db 537 TTGGGATAAAGATTGAAACTGTCCGCTTCTCCACTGCTACAAAAGGGAGTTGATCGGG 596
 QY 233 TGTGATTGACCAACCCCTGTTCTTGGCCAAAGTCTGGGCAAGACCGCTCCAAGCTGT 292
 Db 597 TTTTGTGGATCACCTATGTTTCTGAGAAGGTTTGGGGGAAACAGGAGGAAATAT 656
 QY 293 ACGGCCCCCGCTCCGGCGCTGACTACTGACAAACACAAAGCGCTTCGCCCTGTTCTGCA 352
 Db 657 ATGGTCTCTGTACAGGAACAGATTATGACAGCAATCAACTAAGATTACGCCCTTTGTGCC 716
 QY 353 AGGCGCTATTGAGGCTGCCCGGTGCTGCCCTTCGGCCCCCGGAGGACTGG----- 405
 Db 717 TGGCAGCTCTGGAAGCTCCAAGACTTCTAAATCTCAACACAGCAAAATACTATCTGGAC 776
 QY 406 -----GTCCTGTCGGCCCAAGCACTGGCACTCCGCCCTGGTCCCGTCC 448
 Db 777 CATATGGAGATGATGTTGTTTATTTGCCCACGATTGGCACTCTGCTCTACTGCCCTGCT 836
 QY 449 TGTGGAAGGACGAGTACCAGCCCAAGGGCCAGTTCCACCAAGGCCAAGTCGGTCTGGCTA 508
 Db 837 ACTTGAAGAACTATGTACCAATCACATGATGTTTACATGAATGCTAAGGTTGCAATTGCA 896
 QY 509 TCCCAACATCGCTTCCAGGCGCGCATGTGGAGGAGGCTTTCAGAGCACAGAGCTGC 568
 Db 897 TTCATAATATGCTTACCAGGGCCGATTTGCCCTTTTCGGACTTGAACCTCTTAATCTCC 956
 QY 569 CCCAGGCGCCCTTTCACAAGCTGGCTTCTCGAGCGGTATGCCAAGGTTTACTACTGAGG 628
 Db 957 CCAATAAATTTAAATCTTCATTGATTTCATGATGATATGACAACT----- 1006
 QY 629 CCACCCCATGAGGAGGACGAGAGCCGCCCTGACGGGAAGACCTTACAAGAGATCA 688
 Db 1007 -----GTGAAGGAAGGAAATAA 1025
 QY 689 ACTGCTGAAGGTGGCATATTCGCCGCGCAAGCTGTGACTGTGTGCTGCCCACTAG 748
 Db 1026 ATTGATGAAGGCTGGATAATAGAAATGTAGTGTGCTTGACCTGAGCCCATATATG 1085
 QY 749 CGACCGAGATCGCTGCCGATGCCGCGGTGTGAGCTGGACACCGCTCATCCGCGCA 808
 Db 1086 CCAAGAGCTTCTCAGGGGTAGAGAAGGGTGTGAGTTGGCAATATCTCTCGCATCA 1145
 QY 809 AGGGCATTGAAGGCAATGTGAACGGCATGGACATTGAGAGTGGAAACCCCAAGACGACA 868
 Db 1146 AAACCATCTGTGGAATAGTAATGGATGGACACCGAGTGGAAATCCATTAAAGACA 1205
 QY 869 AGTTCTCTGTCGCGCTACGACACAGACAGCTGTACGCCGCAAGCGCGCCCAAGG 928
 Db 1206 ATATATTCTTACAACATACGATGCAACAACTGTATTTGATGCAAAACCTCTCTGTAAAG 1265
 QY 929 AGGCCCTGCAAGCCGAGCTGGCCCTGCTGTGAGACCCCAACCCGCCCTGTTGCGCTTCA 988
 Db 1266 AAGCTTTCAGCTGAGTGGCTGCCCTGTTTAAACAAAACAAGCTTGTTTGGCCTTG 1325
 QY 989 TCGGCCGCTGGAGGACGAGAGGTGTGGACATCATCTGCCGCCCTGCCCAAGATCC 1048
 Db 1326 TTGAAGACTAGATGAGCAAGAGGCTCAGACATTTCTAGTTCAGCAATTTCCAGAACTC 1385
 QY 1049 TGGCCACCCCAAGGTGAGTCCCATCTCGGTACCGCAAGCCCGCTTACGAGAAGC 1108
 Db 1386 T---TTGTGAGAAATGTTCAAGTGTAGTACTTGGCACTGGCGAAGAAAGTTGGAGATG 1442
 QY 1109 TGGTGAACGCCATCGCCACCAAGTACAAGGCGCGCCCAAGGGCGGTGTCAGTCTCTCGG 1168
 Db 1443 AACITACATTACTTGAAGGAATGTTTCCAGACAAATTCAGACCATCTCAAAATCAAG 1502

XX Wheat starch synthase clone wSIIIB cDNA.
 XX Wheat; starch synthase; SSII; SSIII; starch content; starch synthesis;
 KW food product; adhesive; ss.
 XX Triticum aestivum.
 OS WO200066745-A1.
 PN 09-NOV-2000.
 PD 28-APR-2000; 2000WO-AU00385.
 XX 29-APR-1999; 99AU-0000052.
 XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
 PA (GOOD-) GOODMAN FIELDER LTD.
 PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
 XX Morell M, Li Z, Rahman S, Appels R;
 PI WPI; 2000-647602/62.
 DR Nucleic acid molecules encoding wheat starch synthase (WST)-I and
 PT WST-II, useful in modifying plant starch content and/or composition -
 XX Disclosure; Fig 2; 21pp; English.
 XX The present invention relates to novel protein and coding sequences from
 CC wheat. The proteins are wheat starch synthases, designated SSII and
 CC SSIII. These can be used in the modification of plant starch content or
 CC composition, and to screen plants to identify mutations which affect
 CC starch content and composition. The starch can then be used in food
 CC products, such as flour, and in films, coatings, adhesives, building
 CC materials and packaging materials.
 XX
 XX Sequence 2107 BP; 452 A; 566 C; 664 G; 425 T; 0 other;
 SQ
 Query Match 13.4%; Score 214; DB 21; Length 2107;
 Best Local Similarity 55.8%; Pred. No. 3.9e-25;
 Matches 478; Conservative 0; Mismatches 360; Indels 18; Gaps 3;
 QY 687 CAACTGGCTGAAGGTGGCATTATCCGCCGACAGAGTGGTGTGACGTGTCGCCCACTA 746
 DB 933 CAACTACTCGCGCGCGCTGAAGATGGCGACAGGTGTGCTGGTGGACGCCGGGTA 992
 QY 747 CGGACCGAGATCGTGGCCATCCGCGCGGTGTGGAGCTGGACCGATCC---G 803
 DB 993 CTTGTGGAGCTGAAGACGTGGAGGCGGCTGGGGCTTCACGACATCATACGGCAGAA 1052
 QY 804 CGCCAAGGGCATTGAGGGCATTGTGAACGCGCATGGACATTGAGAGTGAACCCCAAGAC 863
 DB 1053 CGACTGGAACACCGCGGCATCTCAACGCGCATCGACAACTGAGTGAACCCCGAGGT 1112
 QY 864 CGACAGTTCCTGTGCGCCCTACGACCAAGACAGGTGT-----ACGCCG 911
 DB 1113 GGACGCCCACTCAAGTGGCGGCTACACCAACTTCTCGCTGAGGACGCTGGCACTCGG 1172
 QY 912 CAAAGCGCGCCCAAGGAGCGCTGAGCGCGGCTGCTGCTGGACCCCAACCGC 971
 DB 1173 CAAAGCGGCTGAAGAGGCGCTGACGCGGAGCTGGGCTCGAGTCCGCGCGACGT 1232
 QY 972 CCGCTCTTCGCTTCAATCGCGCGCTGGAGGAGCAGAGGGGTGTGACATCATPCCTGGC 1031
 DB 1233 GCGCTCTCGCTTCAATCGCGCGCTGGAGGAGCAGAGGGCGTGGAGATCATCGCGGA 1292
 QY 1032 CGCCCTGCCAAGATCTGCGCACCCCAAGTGCAGATCGCATCTCTGGTACCGCA 1091
 DB 1293 CGCATGCGCTGGATCGT---AGCCAGAGCTGCACTGCTGATGTGGCGACCGGG 1349
 QY 1092 GCGCGCTAGGAGCTGTGACGCGCATCGGCAACCAAGTACAAAGGCGCGCAAGG 1151

DB 1350 CCACGACCTGGAGAGCATGCTGAGCAGCTTCGAGCGGGAGCACCAGCAAGTGGCGG 1409
 QY 1152 CBTGTCAAAGTTCTCGCGCCCTTGGCGACATGTCACCGCGGCGCGGACTTCATGCT 1211
 DB 1410 GTGGGTGGGTTCTCGTGGCTGCGCACCGGATCACGGGGGCGGACGCGCTCT 1469
 QY 1212 GTGCGCTCGCTTCGAGCCCGTGGCGGCTGATCGAGTGCAGCGCATGCACTACGGTAC 1271
 DB 1470 CATGCCCTCCCGTTCTGCTGCGGCTGAACCAAGCTTACGCCATGGCCTACGGCAC 1529
 QY 1272 CGTGGCCGTGTAGCTCCACCGCGGCTGTGTCACCGTCAAGGAGGCGCTCACCGG 1331
 DB 1530 CGTCCCGCTGTCGACCGCTGCGCGGCTCAGGGACACGCTGCGCGCTTCACCCCTT 1589
 QY 1332 CTTCCACATGGCGCCCTGAACCCCGACAAAGCTGGAGGTGACCGCGACGCCCTGGC 1391
 DB 1590 CAACCACTCGGGCTCGGTGGAGCTTCGACCGCGCGAGGCGCAACAAGCTGATCGAGG 1649
 QY 1392 CGCCACGTCGCGGCTGCGACGAGGTGTGCGGCGCGGCTACCCCGAGATGTGGC 1451
 DB 1650 GCTCGGCACTGCTCCGACCTACCGAGACTTCAAGGAGAGCTGGAGGCGCTCCAGGA 1709
 QY 1452 CAACTGATCAGCCAGGACCTGCTGTCCTCAAGCCCGCCAGAGTGGAGGCGCTGCT 1511
 DB 1710 GCGCGCATGTCGAGGACTTCACTGGAGACGCGCGCAAGCTCTACGAGGACGCTCT 1769
 QY 1512 GGAGGAGGTGGTGTAC 1527
 DB 1770 CGTCAAGGCCAAGTAC 1785
 RESULT 15
 AAC86431
 ID AAC86431 standard; DNA; 9024 BP.
 XX
 AC AAC86431;
 XX
 DT 01-MAR-2001 (first entry)
 XX
 DE Wheat SSII gene SEQ ID NO: 37.
 XX
 KW Wheat; starch synthase; SSII; SSIII; starch content; starch synthesis;
 KW food product; adhesive; ds.
 XX
 OS Triticum tauschii.
 XX
 PN WO200066745-A1.
 PD 09-NOV-2000.
 XX
 PF 28-APR-2000; 2000WO-AU00385.
 XX
 PR 29-APR-1999; 99AU-0000052.
 XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
 PA (GOOD-) GOODMAN FIELDER LTD.
 PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
 XX
 PI Morell M, Li Z, Rahman S, Appels R;
 DR WPI; 2000-647602/62.
 XX
 PT Nucleic acid molecules encoding wheat starch synthase (WST)-I and
 PT WST-II, useful in modifying plant starch content and/or composition -
 XX
 PS Claim 1; Page 195-200; 21pp; English.
 XX
 CC The present invention relates to novel protein and coding sequences from
 CC wheat. The proteins are wheat starch synthases, designated SSII and
 CC SSIII. These can be used in the modification of plant starch content or
 CC composition, and to screen plants to identify mutations which affect
 CC starch content and composition. The starch can then be used in food
 CC products, such as flour, and in films, coatings, adhesives, building

CC materials and packaging materials.

XX SQ

SQ Sequence 9024 BP; 2228 A; 2386 C; 2216 G; 2193 T; 1 other;

Query Match 13.3%; Score 212.4; DB 21; Length 9024;
Best Local Similarity 55.7%; Pred. No. 6.5e-25;
Matches 477; Conservative 0; Mismatches 361; Indels 18; Gaps 3;

QY	687	CAAC	TGCTGAAGGGTGGCAN	TATGCCCGCAGAGCTGGT	GACTGTTGTCGCCCACTA	746
Db	7392	CAACT	TTCCGCGCGGCTGAAGAT	GGCGACAGGTTGTCGTGGT	GAGCCCGGGA	7451
QY	747	CGCAG	CCAGATCGCTGCCGAT	CCCGCGCGTGTGGAGCT	GGACACCGTCATCC---	803
Db	7452	CCTGG	GAGCTGAGACGGTGGAG	GGCGGCTGGGGGCTTCACGACAT	CATACGSGCAA	7511
QY	804	CGC	CAAGGCATTGAGGCATT	GTGAACGGCATGGACATT	GAGGAGTGGAAACCCCAAGAC	863
Db	7512	CGACT	GGAAGACCCGGGCATCT	CAACGGCATCGACAACT	ATGGAATGGAAACCCCGAGT	7571
QY	864	CGACA	AGTTCTGTCTGCGCC	TACACCAAGACAGCTCT-----	-AGCGCGG	911
Db	7572	GGAC	GGCCACTCAAGTGGAGGG	TACACCAACTTCTCCCTGAGGAG	CGCTGGAATCCGG	7631
QY	912	CAAG	CGCGCCCAAGAGGCGCT	GCAGCGCAGCTGGCGT	GCTGTGGACCCACACGC	971
Db	7632	CAAG	CGCAGTGCAGGAGGCGCT	GCAGCGCAGCTGGCGT	GCAGGTCGCGCGCAGCT	7691
QY	972	CCCG	CTGTGCGCTTCATCGCG	CGCTGGAGGACAGAGGT	GTGGACATCATCTCTGGC	1031
Db	7692	GCG	CTGTGCGCTTCATCGCG	CGCTGGAGGACAGAGGT	GTGGAGATCATCGCGA	7751
QY	1032	CGCC	TGCCCAAGATCTCGCC	ACCCCAAGGTGAGATGCCAT	CTGGGTACCGGCAA	1091
Db	7752	CGCA	TGCGCTGATCGTG-	--AGCCAGCAGCTGAGTGT	GATGCTGGGCAACCGGCG	7808
QY	1092	GGCG	CGTACGAGAGCTGGTGA	AGCCCATGGCACCAAGTACA	AGGCGCGCGCAAGG	1151
Db	7809	CCAC	AGCTGGAGAGCATCT	CGCGCACTTCGAGCGGAG	CACACACAGGTGCGCG	7868
QY	1152	CGT	GTCAAGTTCTCGGCG	CCCTGCGGCATGTCAC	CGCGCGGCACATTCATGCT	1211
Db	7869	GTGG	TGGGTCTCGTGGCG	CTCGGCAACGGATCAGCG	GGGGCGACGCGTCT	7928
QY	1212	GGT	CGCTCGCGTTGAGC	CCCTGCGGCTGATCAG	CTGCACGCCATGCACATCGGTAC	1271
Db	7929	CAT	CGCTCCCGTTGAGCG	CTGCGGGCTGAACAG	CTCTACGCCATGCGCTACGCGAC	7988
QY	1272	CGT	GCCGTGTAGCTCCAC	CGGCGCTGGTGACAC	GTCAACGTCAAGAGGGCGTCA	1331
Db	7989	CGT	CCCGTGTGACGCG	CTGGGGCTCAGGAG	CAACGCTGCGCGCTTCGACCCCTT	8048
QY	1332	CTT	CCACATGGCGCCCTGA	ACCCGACAGCTGACGAG	CTCACCGCAGCGCTTGGC	1391
Db	8049	CAAC	CACTCGGGCTCGGTGG	AGCTTCGACCGCGCAG	CGGCACAAGCTGATCAGGC	8108
QY	1392	CGC	ACCGTGGCCGTGCC	AGGAGTGTTCGGGGCGGCG	CTACCCCGAGATGGTGGC	1451
Db	8109	GCT	CGGCACTGCCTCCG	CACTACCGAGCTTCAAG	SAGAGTGGAGGGCCCTCCAGGA	8168
QY	1452	CAAC	TGCATCAGCCAGGAC	CTCTGTGTCCAAAGCG	CGCCAGAAATGGGAGGCGTGCCT	1511
Db	8169	GC	CGCATGCGAGGACT	TCAGCTGGAG	ACGCGCGCAAGCTCTACGAGGACGTCT	8228
QY	1512	GGAG	GAGTGTGTAC	1527		
Db	8229	CGT	CAAGGCCAAGTAC	8244		

Search completed: June 3, 2003, 10:09:38
Job time : 279.514 secs

This Page Blank (uspto)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2003, 09:33:29 ; Search time 1554.22 Seconds
(without alignments)
16599.542 Million cell updates/sec

Title: US-09-980-771A-8
Perfect score: 1593
Sequence: 1 gcgtgacatcgtgatggt.....ccgtgcccgaagaatcccc 1593

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_mam.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	469.8	29.5	473	10	AV640741
2	457.8	28.7	461	10	AV629488
3	454.6	28.5	2147	11	AY109531
4	444.4	27.9	458	10	AV622787
5	400.8	25.2	436	10	AV629117
6	395.6	24.8	511	10	AV394077

7	365.2	22.9	475	10	AW757933
8	351.2	22.0	799	12	BF863935
9	337.8	21.2	418	10	AV644765
10	295.8	18.6	506	10	AV643281
11	278.4	17.5	702	14	BQ246353
12	272.8	17.1	560	10	AV628313
13	267	16.8	393	10	AV631115
14	251.8	15.8	552	10	BE024926
15	250.2	15.7	517	10	AV641583
16	245.6	15.4	726	14	BQ804991
17	244.2	15.3	532	10	AV641989
18	242	15.2	585	14	BQ246366
19	235.8	14.8	513	10	AV631004
20	235.8	14.8	552	10	AV640546
21	234.8	14.7	524	10	AV641724
22	234.2	14.7	500	10	AV642869
23	233	14.6	502	14	BQ814687
24	231.8	14.6	510	10	AV393307
25	230.4	14.5	761	12	BG366176
26	223.8	14.0	540	10	AV642722
27	222.2	13.9	1545	11	AY112448
28	216.6	13.6	773	12	BG351175
29	212	13.3	572	14	BQ245812
30	211.8	13.3	486	10	AV642834
31	211.8	13.3	504	10	AV629198
32	210.8	13.2	502	10	AV644517
33	210.2	13.2	511	10	BE423250
34	209.2	13.1	696	12	BG351920
35	208.8	13.1	495	10	AV642565
36	208.2	13.1	543	14	BQ245428
37	207	13.0	641	14	BQ247154
38	206.8	13.0	522	10	BE423625
39	206.4	13.0	729	12	BG599615
40	200.8	12.6	493	10	AV631838
41	200.2	12.6	2498	11	AY109714
42	200	12.6	631	13	BF531619
43	199.8	12.5	492	10	AV629891
44	198.8	12.5	484	10	AV629266
45	198.8	12.5	489	10	AV629190

ALIGNMENTS

RESULT 1
AV640741
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT
FEATURES
source

AV640741 473 bp mRNA linear EST 15-DEC-2000
AV640741 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
CDNA clone HCL021f06_r 5', mRNA sequence.
AV640741.1 GI:10784069
EST.
Chlamydomonas reinhardtii.
Chlamydomonas reinhardtii
Eukaryota, Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
1 (bases 1 to 473)
Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K., Nakamura,Y. and Tabata,S.
Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii
DNA Res. 7 (5), 305-307 (2000)
20539644
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizuekazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1. 473
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"

```
/clone="HCL021f06_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/notes="vector: pBluescriptII SK-; Site.1: EcoRI; Site.2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT      76 a 162 c 146 g 89 t
ORIGIN
Query Match      29.5%; Score 469.8; DB 10; Length 473;
Best Local Similarity 99.6%; Pred. No. 1.1e-77;
Matches 471; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 17 TGGTGTGCTGAGTGGCCCTTGGTCCAAAGACGGCGGCTGGGCGATGTGACTGGTG 76
DB 1 TGGTGTGCTGAGTGGCCCTTGGTCCAAAGACGGCGGCTGGGCGATGTGACTGGTG 60
QY 77 GCCTGCTATTTAGAGTGTGAAGCGGCGCCACCGCGTCATGACCATGCCCCCTCGCTACG 136
DB 61 GCCTGCTATTTAGAGTGTGAAGCGGCGCCACCGCGTCATGACCATGCCCCCTCGCTACG 120
QY 137 ACCAGTAGCTGAGCTGGGACACCTCGGTGGTGGACATCATGGCGGAGAGGTCC 196
DB 121 ACCAGTAGCTGAGCTGGGACACCTCGGTGGTGGACATCATGGCGGAGAGGTCC 180
QY 197 GCTACTTCCATCCATCAAGAAGGCGGTGCACCGCGTGTGATTGACCAACCCCTGGTTCC 256
DB 181 GCTACTTCCATCCATCAAGAAGGCGGTGCACCGCGTGTGATTGACCAACCCCTGGTTCC 240
QY 257 TGGCCAGGTCTGGGCGAAGACCGGCTCCAAAGCTGTAGGCGCCCGCTCCGGCGTGTACT 316
DB 241 TGGCCAGGTCTGGGCGAAGACCGGCTCCAAAGCTGTAGGCGCCCGCTCCGGCGTGTACT 300
QY 317 ACCTGGACCAACCAAGCGCTTCGCCCTTTCTGCAAGCGCGCTATTGAGGCTGCCCGCG 376
DB 301 ACCTGGACCAACCAAGCGCTTCGCCCTTTCTGCAAGCGCGCTATTGAGGCTGCCCGCG 360
QY 377 TGCTGCCCTTCGGCCCGCGGAGGACTGCGTCTTGGTGGCCAAAGCTGGCACTCCGCC 436
DB 361 TGCTGCCCTTCGGCCCGCGGAGGACTGCGTCTTGGTGGCCAAAGCTGGCACTCCGCC 420
QY 437 TGCTGCCCTTCGCTGAAGACAGTACCAAGCCCAAGGCCAGTTCACCAAG 489
DB 421 TGCTGCCCTTCGCTGAAGACAGTACCAAGCCCAAGGCCAGTTCACCAAG 473

RESULT 2
AV629488 461 bp mRNA linear EST 15-DEC-2000
LOCUS
DEFINITION Chlamydomonas reinhardtii 5% to 0.04% CO2 Chlamydomonas
reinhartii cDNA clone LCL059f09_r 5', mRNA sequence.
ACCESSION AV629488
VERSION AV629488.1 GI:10792122
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE
AUTHORS Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
1 (bases 1 to 461)
Asamizu, E., Miura, K., Kuchio, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,
Nakamura, Y. and Tabata, S.
TITLE Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
JOURNAL DNA Res. 7 (5), 305-307 (2000)
MEDLINE 20539644
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizukazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
1. 461
/organism="Chlamydomonas reinhardtii"

/strain="C9"
/db_xref="taxon:3055"
/clone="LCL059f09_r"
/clone_lib="Chlamydomonas reinhardtii 5% to 0.04% CO2"
/notes="vector: pBluescriptII SK-; Site.1: EcoRI; Site.2:
XhoI; The cDNA library was constructed from cells cultured
in a carbon stress acclimated condition in which carbon
dioxide concentration in the bubbling gas was changed from
5% to 0.04%"
BASE COUNT      73 a 156 c 143 g 89 t
ORIGIN
Query Match      28.7%; Score 457.8; DB 10; Length 461;
Best Local Similarity 99.6%; Pred. No. 1.9e-75;
Matches 459; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 8 ACATCGTATGTTGCTGCTGAGTCCGCCCTTGGTCCAAAGACGGCGGCTGGGCGATG 67
DB 1 ACATCGTATGTTGCTGCTGAGTCCGCCCTTGGTCCAAAGACGGCGGCTGGGCGATG 60
QY 68 TGACTGGTGGCTGCTGCTATTTAGCTGGTCAAGCGGCGCCACCGCGTCATGACCAITGCC 127
DB 61 TGACTGGTGGCTGCTGCTATTTAGCTGGTCAAGCGGCGCCACCGCGTCATGACCAITGCC 120
QY 128 CTCGTAGCACAGTACGCTGACGCTGGGACACCTCGTGGTGGTGGACATCATGGGCG 187
DB 121 CTCGTAGCACAGTACGCTGACGCTGGGACACCTCGTGGTGGTGGACATCATGGGCG 180
QY 188 AGAAGTCCGCTACTTCCACTCCATCAAGAGGGGCTGCACCGCGTGTGGATTGACCAAC 247
DB 181 AGAAGTCCGCTACTTCCACTCCATCAAGAGGGGCTGCACCGCGTGTGGATTGACCAAC 240
QY 248 COTGTTCTCTGSCCAAGGTCTGGGCAAGACCGGCTCCAAAGCTGTACGCGCCCGCTCG 307
DB 241 COTGTTCTCTGSCCAAGGTCTGGGCAAGACCGGCTCCAAAGCTGTACGCGCCCGCTCG 300
QY 308 GGCCTGACTACTCTGCACACACACCGCTTCGCCCTTCTGCAAGGCGCTATTGAGG 367
DB 301 GGCCTGACTACTCTGCACACACACCGCTTCGCCCTTCTGCAAGGCGCTATTGAGG 360
QY 368 CTGCCCGGCTGTCGCCCTTCGCCCGCGGAGGACTGCGTCTTCGTGGCCAAAGCTGCG 427
DB 361 CTGCCCGGCTGTCGCCCTTCGCCCGCGGAGGACTGCGTCTTCGTGGCCAAAGCTGCG 420
QY 428 ACTCGGCCCTGTCGCCCGCTGCTGCTGAAGGACGAGTACCAG 468
DB 421 ACTCGGCCCTGTCGCCCGCTGCTGCTGAAGGACGAGTACCAG 461

RESULT 3
AY109531 2147 bp mRNA linear HTC 25-MAY-2002
LOCUS
DEFINITION Zea mays CL1198_1 mRNA sequence.
ACCESSION AY109531
VERSION AY109531.1 GI:21213285
KEYWORDS HTC.
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 2147)
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 2147)
AUTHORS Coe, E.C.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
FEATURES
Location/Qualifiers
```


Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1. 458
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone_lib="Chlamydomonas reinhardtii 5% to 0.04% CO2"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from cells cultured
in a carbon stress acclimatized condition in which carbon
dioxide concentration in the bubbling gas was changed from
5% to 0.04%"

BASE COUNT 91 a 147 c 146 g 74 t

Query Match 27.9%; Score 444.4; DB 10; Length 458;
Best Local Similarity 99.6%; Pred. No. 5.8e-73;
Matches 456; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 365 AGGCTGCCCGCTGCTGCCCTTCCGCGCGGAGGAGTCCGCTCTCGTGGCCCAAGGACT 424
|||||
Db 1 AGGCTGCCCGCTGCTGCCCTTCCGCGCGGAGGAGTCCGCTCTCGTGGCCCAAGGACT 60
QY 425 GGCACCTCCGCCCTGGTGGCCCTCTGCTGAAGGACGAGTACCAGCCCAAGGGCCAGTTCA 484
|||||
Db 61 GGCACCTCCGCCCTGGTGGCCCTCTGCTGAAGGACGAGTACCAGCCCAAGGGCCAGTTCA 120
QY 485 CCAAGGCCAAGTGGTGGTGTATC-CACACATCGCCCTTCCAGGGCCGCGATGGGGAG 543
|||||
Db 121 CCAAGGCCAAGTGGTGGTGTATCGCACAAATCGCCCTTCCAGGGCCGCGATGGGGAG 180
QY 544 GAGGCTTTCAAGGACACGAGTGGCCCGAGCGCGCTTTGACAAGCTGGCTTCTCGGAC 603
|||||
Db 181 GAGGCTTTCAAGGACACGAGTGGCCCGAGCGCGCTTTGACAAGCTGGCTTCTCGGAC 240
QY 604 GGCTATGCCAAGTTTACACTGAGGCCACCCCATGGAGGAGGACGAGAGCCCGCGCTG 663
|||||
Db 241 GGCTATGCCAAGTTTACACTGAGGCCACCCCATGGAGGAGGACGAGAGCCCGCGCTG 300
QY 664 ACGGAAAGACCTACAAGAATFCACTGGCTGAAGGGTGGCATTTATCGCGCCGACAG 723
|||||
Db 301 ACGGAAAGACCTACAAGAATFCACTGGCTGAAGGGTGGCATTTATCGCGCCGACAG 360
QY 724 CTGGTCACTGTGCGCCCACTACGGACCGAGATCGCTGCCATGCCGCGCGGCTGTG 783
|||||
Db 361 CTGGTCACTGTGCGCCCACTACGGACCGAGATCGCTGCCATGCCGCGCGGCTGTG 420
QY 784 GAGCTGGACACCGCTATCCCGGCCCAAGGGCAITGAGGG 821
|||||
Db 421 GAGCTGGACACCGCTATCCCGGCCCAAGGGCAITGAGGG 458

RESULT 5
AV629117
LOCUS
DEFINITION AV629117 Chlamydomonas reinhardtii 5% to 0.04% CO2 Chlamydomonas
reinhardtii cDNA clone LCL052d05_r 5', mRNA sequence.
ACCESSION AV629117
VERSION AV629117.1 GI:10791751
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaeae; Chlamydomonas.
REFERENCE 1 (bases 1 to 436)
AUTHORS Asamizu, E., Miura, K., Kuchio, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,
Nakamura, Y., and Tabata, S.
TITLE Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
JOURNAL DNA Res. 7 (5), 305-307 (2000)

205399644

Medline
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1. 436
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone_lib="Chlamydomonas reinhardtii 5% to 0.04% CO2"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from cells cultured
in a carbon stress acclimatized condition in which carbon
dioxide concentration in the bubbling gas was changed from
5% to 0.04%"

FEATURES
source

BASE COUNT 66 a 145 c 138 g 87 t

Query Match 25.2%; Score 400.8; DB 10; Length 436;
Best Local Similarity 99.5%; Pred. No. 7.4e-65;
Matches 402; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGCTGGACATCGTATGTTGCTGCTGAGTGGCCCTTGGTCCAAAGACCGCGGCTG 60
|||||
Db 33 GGCTGGACATCGTATGTTGCTGCTGAGTGGCCCTTGGTCCAAAGACCGCGGCTG 92
QY 61 GCGATGTGACTGTGCTGCTGCTATTGAGTGTCAAGCGCGCCACCGCGTCATGACC 120
|||||
Db 93 GCGATGTGACTGTGCTGCTGCTATTGAGTGTCAAGCGCGCCACCGCGTCATGACC 152
QY 121 ATTGCCCTCGCTACGACCACTAGTACGCTGACGCTGGGACACCTCGGTGCTGGACATC 180
|||||
Db 153 ATTGCCCTCGCTACGACCACTAGTACGCTGACGCTGGGACACCTCGGTGCTGGACATC 212
QY 181 ATGGGGAGAAGTCCGCTACTTCCACTCCATCAAGAGGGCGTGCACCGGCTGGATT 240
|||||
Db 213 ATGGGGAGAAGTCCGCTACTTCCACTCCATCAAGAGGGCGTGCACCGGCTGGATT 272
QY 241 GACCACCCCTGTTCTTGGCCAAAGTCTTGGGCAAGACCGCGTCCAAAGCTGTACGGCCCC 300
|||||
Db 273 GACCACCCCTGTTCTTGGCCAAAGTCTTGGGCAAGACCGCGTCCAAAGCTGTACGGCCCC 332
QY 301 CGCTCCGCGCTGACTACCTGGAACACCAAGCGCTTCGCCCTGTCTGCAAGGCCGCT 360
|||||
Db 333 CGCTCCGCGCTGACTACCTGGAACACCAAGCGCTTCGCCCTGTCTGCAAGGCCGCT 392
QY 361 ATTGAGGCTCCCGCGTGTGCTGCCCTTCCGCGCGCGGAGGACTG 404
|||||
Db 393 ATTGAGGCTCCCGCGTGTGCTGCCCTTCCGCGCGCGGAGGACTG 436

RESULT 6
AV394077
LOCUS
DEFINITION AV394077 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
cDNA clone CL18c08_r 5', mRNA sequence.
ACCESSION AV394077
VERSION AV394077.1 GI:6548293
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaeae; Chlamydomonas.
REFERENCE 1 (bases 1 to 511)
AUTHORS Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H., and Tabata, S.
TITLE A large scale structural analysis of cDNAs in a unicellular green
alga, Chlamydomonas reinhardtii. I. Generation of 3433
non-redundant expressed sequence tags
JOURNAL DNA Res. 6 (6), 369-373 (1999)
MEDLINE 20152988

COMMENT

Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/eu/plant/.

FEATURES

source

1. .511
Location/Qualifiers
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone="CL18c08_r"
/clone_lib="Chlamydomonas reinhardtii C9"
/dev_stage="photoautotrophic growth"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 80 a 170 c 159 g 101 t 1 others

Query Match 24.8%; Score 395.6; DB 10; Length 511;
Best Local Similarity 98.8%; Pred. No. 7e-64;
Matches 398; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCGCTGGACATCGTATGTTGCTGTGAGTGGCGCCCTTGTCCTCAAGACGGCGGCTG 60
DB 109 GCGCTGGACATCGTATGTTGCTGTGAGTGGCGCCCTTGTCCTCAAGACGGCGGCTG 168
QY 61 GCGGATGTGACTGGTGGCTGCTATTGAGTGGTCAAGCGGGCCACCGCGTATGACC 120
DB 169 GCGGATGTGACTGGTGGCTGCTATTGAGTGGTCAAGCGGGCCACCGCGTATGACC 228
QY 121 ATTGCCCTCGCTACGACAGTACGCTGACGCTGGGACACCTCGTGTGCTGGACATC 180
DB 229 ATTGCCCTCGCTACGACAGTACGCTGACGCTGGGACACCTCGTGTGCTGGACATC 288
QY 181 ATGGGGGAGAAGTGGCTGCTTCCACTCCATCAAGAGGGGGTGCACCGCGTGTGGATT 240
DB 289 ATGGGGGAGAAGTGGCTGCTTCCACTCCATCAAGAGGGGGTGCACCGCGTGTGGATT 348
QY 241 GACCACTCGCTGCTGCGCAAGTCTGGGCAAGACCGGCTTCAAGTGTACGCGCCC 300
DB 349 GACCACTCGCTGCTGCGCAAGTCTGGGCAAGACCGGCTTCAAGTGTACGCGCCC 408
QY 301 CGCTCGGCGCTGACTTACCTGGACACCAACAGCGCTTCGCCCTGTTCGCAAGCGCGT 360
DB 409 CGCTCGGCGCTGACTTACCTGGACACCAACAGCGCTTCGCCCTGTTCGCAAGCGCGT 468
QY 361 ATTGAGCTGCGCGCTGCTGCTGCTGCGCGCGCGAGGACT 403
DB 469 ATAGAGGCTGCGCGCTGCTGCTGCGCGCGCGAGGACT 511

RESULT 7

AW757933

LOCUS

DEFINITION 874004E12.y1 C. reinhardtii CC-1690, Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.

ACCESSION

AW757933

VERSION

AW757933.1

KEYWORDS

EST.

SOURCE

Chlamydomonas reinhardtii.

Chlamydomonas reinhardtii

Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

Chlamydomonadaceae; Chlamydomonas.

1 (bases 1 to 475)

Grossman, A., Davies, J., Federspiel, N., Harris, E., Lefebvre, P.,

McDermott, J.P., Silflow, C., Stern, D., and Surzycki, R.

Analyses of the Chlamydomonas reinhardtii Genome: A Model,

Unicellular System for Analyzing Gene Function and Regulation in

Vascular Plants: project phase 2

Unpublished (2000)

Contact: Elizabeth H. Harris

DCMB Box 91000

Duke University

Durham, NC 27708-1000, USA
Tel: 919 613 8164
Fax: 919 613 8177
Email: chlamy@duke.edu.

FEATURES

source

1. .475
Location/Qualifiers
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Lambda Zap II"
/note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2: XhoI; This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in YAP (acetate-containing) medium in the light, YAP medium in the dark, HS (minimal) medium in ambient levels of CO₂ and HS medium bubbled with 5% CO₂. POLYA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with EXAssist (Stratagene) phage."

BASE COUNT 74 a 160 c 143 g 98 t

Query Match 22.9%; Score 365.2; DB 10; Length 475;
Best Local Similarity 97.9%; Pred. No. 3.1e-58;
Matches 370; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 CGCTGACATCGTATGTTGCTGTGAGTGGCGCCCTTGTCCTCAAGACGGCGGCTG 61
DB 98 CGCTGACATCGTATGTTGCTGTGAGTGGCGCCCTTGTCCTCAAGACGGCGGCTG 157
QY 62 GCATGTGACTGTGGCTGCTTATTGAGTGGTCAAGCGGGCCACCGGTCATGACCA 121
DB 158 GCATGTGACTGTGGCTGCTTATTGAGTGGTCAAGCGGGCCACCGGTCATGACCA 217
QY 122 TTGCCCTCGCTACGACAGTACGCTGACGCTGGACACCTCGGTGGTGGACATCA 181
DB 218 TTGCCCTCGCTACGACAGTACGCTGACGCTGGACACCTCGGTGGTGGACATCA 277
QY 182 TGGCGGAGAGTTCGCTACTTCCATCCATCAAGAGGGCGTGCACCGGTGTGGATT 241
DB 278 TGGCGGAGAGTTCGCTACTTCCATCCATCAAGAGGGCGTGCACCGGTGTGGATT 337
QY 242 ACCACCCCTGGTTCCTGGGCAAGTCTGGGGCAAGACCGGCTCCAAAGTGTACGGCCCC 301
DB 338 ACCACCCCTGGTTCCTGGGCAAGTCTGGGGCTAGACCGGCTCCAAAGTGTACGGCCCC 397
QY 302 GCTCGGCGCTGACTACCTGGGACACCAAGCGCTTCGCCCTGTTCGCAAGCGCGCTA 361
DB 398 GCTCGGCGCTGACTACCTGGGACACCAAGCGCTTCGCCCTGTTCGCAAGCGCGCTA 457
QY 362 TTGAGGCTGCGCGGTCG 379
DB 458 TTGAGGCTGCGCGGTCG 475

RESULT 8

BF863935

LOCUS

DEFINITION

BF863935

ACCESSION

BF863935

VERSION

BF863935.1

KEYWORDS

EST.

SOURCE

Chlamydomonas reinhardtii.

Chlamydomonas reinhardtii

Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

Chlamydomonadaceae; Chlamydomonas.

1 (bases 1 to 799)

Grossman, A., Davies, J., Federspiel, N., Harris, E., Hauser, C.,

Lefebvre, P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.

TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants; project phase 3

JOURNAL Unpublished (2000)

COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.

FEATURES Location/Qualifiers
source
1. 799
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress condition I, normalized, lambda zap II"
/note="vector: pBluescript II SK-; Site_1: EcoRI; Site_2: XhoI; This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP-N (30 min, 1hr, 4hr), TAP-S (30 min, 1hr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, 1hr, 4hr) and NH4 to NO3 (30min, 1hr, 4hr). PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoRI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."

BASE COUNT 141 a 256 c 208 g 193 t 1 others

ORIGIN

Query Match 22.0%; Score 351.2; DB 12; Length 799;
Best Local Similarity 98.9%; Pred. No. 1.4e-55;
Matches 364; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1155 GGTCAAGTCTTCGGCGCCCTGGCGCACATGCTCACCGCGCGCGACTTCATGCTGT 1214
Db 124 GGTCAAGTCTTCGGCGCCCTGGCGCACATGCTCACCGCGCGCGACTTCATGCTGT 183

Qy 1215 GCCCTCGGCTTCGAGCCCTGGGGCTGATCCAGTGCACGCCATGCACTACGGTACCGT 1274
Db 184 GCCCTCGGCTTCGAGCCCTGGGGCTGATCCAGTGCACGCCATGCACTACGGTACCGT 243

Qy 1275 GCCCTCGGCTTCGAGCCCTGGGGCTGATCCAGTGCACGCCATGCACTACGGTACCGT 1334
Db 244 GCCCTCGGCTTCGAGCCCTGGGGCTGATCCAGTGCACGCCATGCACTACGGTACCGT 303

Qy 1335 CCACATGGCGCCCTGAACCCCGACAAAGTGCACGAGGTGCACGCCGCGCCCTGGCCGC 1394
Db 304 CCACATGGCGCCCTGAACCCCGACAAAGTGCACGAGGTGCACGCCGCGCCCTGGCCGC 363

Qy 1395 CACCTGGCGCTGCCAGGAGGTGTTTCGGGGCGCGCTACCCCGAGATGGTGGCCAA 1454
Db 364 CACCTGGCGCTGCCAGGAGGTGTTTCGGGGCGCGCTACCCCGAGATGGTGGCCAA 423

Qy 1455 CTGCATCAGCAGGACCTGCTCTGTCCTCAAGCCGCCAGAGTGGAGGGCC-TGCTGG 1513
Db 424 CTGCATCAGCAGGACCTGCTCTGTCCTCAAGCCGCCAGAGTGGAGGGCC-TGCTGG 483

Qy 1514 AGGAGGTG 1521
Db 484 AGGAGGTG 491

RESULT 9
AV644765 418 bp mRNA linear EST 15-DEC-2000
LOCUS
DEFINITION AV644765 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
CDNA clone HCL093d12_r 5', mRNA sequence.

ACCESSION AV644765

VERSION AV644765.1 GI:10788093

KEYWORDS EST.

SOURCE Chlamydomonas reinhardtii.

ORGANISM Chlamydomonas reinhardtii

REFERENCE Eukaryota, Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
AUTHORS 1 (bases 1 to 418)
Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K., Nakamura, Y. and Tabata, S.
Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii

TITLE DNA Res. 7 (5), 305-307 (2000)

JOURNAL 20539644

MEDLINE

COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES Location/Qualifiers
source
1. 418
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone_lib="HCL093d12_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/note="vector: pBluescript II SK-; Site_1: EcoRI; Site_2: XhoI; The cDNA library was constructed from cells cultured in a medium with bubbling air containing 5% carbon dioxide"

BASE COUNT 65 a 139 c 130 g 84 t

ORIGIN

Query Match 21.2%; Score 337.8; DB 10; Length 418;
Best Local Similarity 99.4%; Pred. No. 3.8e-53;
Matches 339; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGCCTGGACATCGTATGTTGCTGCTGAGTGCGCCCTTGTTCACAGACGGCGCCG 60
Db 78 GGCCTGGACATCGTATGTTGCTGCTGAGTGCGCCCTTGTTCACAGACGGCGCTG 137

Qy 61 GCGATGTGACTGGTGGCTGCTATTGAGCTGGTCAAGCGGGCCACCGCGTCAATGACC 120
Db 138 GCGATGTGACTGGTGGCTGCTATTGAGCTGGTCAAGCGGGCCACCGCGTCAATGACC 197

Qy 121 ATTGCCCTTCGCTACGACCAAGTACGCTGACGCTGCGACACCTCGGTGGTGGACATC 180
Db 198 ATTGCCCTTCGCTACGACCAAGTACGCTGACGCTGCGACACCTCGGTGGTGGACATC 257

Qy 181 ATGGGGGAGAAAGTTCGCTACTTCCACTCCATCAACAAGGGCGTGCACCGGTGTGATT 240
Db 258 ATGGGGGAGAAAGTTCGCTACTTCCACTCCATCAACAAGGGCGTGCACCGGTGTGATT 317

Qy 241 GACCAACCCCTGGTTCCTGGCCAAAGTCTGGGGCAAGACCGGCTCCCAAGTGTACGGCCCC 300
Db 318 GACCAACCCCTGGTTCCTGGCCAAAGTCTGGGGCAAGACCGGCTCCCAAGTGTACGGCCCC 377

Qy 301 CGCTCCGGCGCTGACTTACCTGGACACCAAGCGCTTCGC 341
Db 378 CGCTCCGGCGCTGACTTACCTGGACACCAAGCGCTTCGC 418

RESULT 10
AV643281 506 bp mRNA linear EST 15-DEC-2000
LOCUS
DEFINITION AV643281 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
CDNA clone HCL066h09_r 5', mRNA sequence.

ACCESSION AV643281

VERSION AV643281

KEYWORDS EST.

SOURCE Chlamydomonas reinhardtii.

ORGANISM Chlamydomonas reinhardtii

DEFINITION Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas

```

Email: scloutier@em.agr.ca
was cloned directionally, not all sequences generated with reverse
primer were from the 5' end (same with forward primer and 3' end).
Average insert size is >1.4 kb
Plate: 012 row: H column: 04
Seq primer: M13 Reverse.
Location/Qualifiers
1. 702
/organism="Triticum aestivum"
/cultivar="Glenlea"
/db_xref="taxon:4565"
/clone="TaEL5012H04R"
/clone_lib="TaE15"
/tissue_type="developing seeds"
/dev_stage="15 days after anthesis"
/lab_host="E. coli DH10B"
/note="vector: pCMV-SPORT6.0 (Invitrogen Technologies);
Site_1: NotI; Site_2: MluI; mRNA obtained from wheat seeds
of cultivar Glenlea 15 days post-anthesis"
143 a 237 q 109 t
BASE COUNT

```

Query Match	17.58;	Score 278.4;	DB 14;	Length 702;
Best Local Similarity	66.5%;	Pred. No. 4.6e-42;		
Matches 399;	Conservative 0;	Mismatches 201;	Indels 0;	Gaps 0;
QY	682	AAGATCAACTGGCTGGAAGGGTGGCATTCGCCGCCGACAAGCTGGTGGACATGTTGGCC	741	
Db	103	AAGATCAACTGGATGAAGCGCGGATCCTCGAGCCGACAAAGTGTCTGACGGTGAAGCCC	162	
QY	742	AACTACCGGACCGAGATCGTCCGATCGCCGCCGGGTGTGGAGCTGGACACCGTCAAC	801	
Db	163	TACTACCGGAGGAGCTCACTCTGGCGAAGCCAGGGGCTCGAGAGCTCGACAACATCATG	222	
QY	802	CGGCCCAAGGSCATTTAGGGCATTTGAACGGCATGGACATTTAGGAGTGAACCCCAAG	861	
Db	223	CGCTCACTGGGATCACCGGATCGTCAACGGCATGTATTACGGAGTGGGAGCCCAACC	282	
QY	862	ACGACAAAGTTCTCTGTGGGCCCTACGACACAGAACAGCGTCTACGCCGGCAAGCGCGCC	921	
Db	283	AAGSACAAGTTCTCGCCGTCAACTAGACATCACACCGGTTGGAGGGGAAGCGCGTG	342	
QY	922	GCAAGAGGCGCCTGCAAGGCGGAGCTGGGCGTGCTGTGGACCCCAACGCCGCCCTGTTC	981	
Db	343	AACAAGAGGCGGTGCAAGGCGGAGTGGGGTGCCGGTGGACCGGAAGTGCCTGGTG	402	
QY	982	GCCTTCATCGGCGCCTGGAGGAGCAGAAGGGTGTGCACATCATCTTGGCGGCCCTGGCC	1041	
Db	403	CGGTTATCGGCAGCTTGGAGGAGCAGAAGGGGCCCAAGCTGATGATTCGGCGCATCCG	462	

[illegible]

```

VERSION AV628313.1 GI:10790947
SOURCE EST.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE Chlamydomonas reinhardtii.
AUTHORS Chlamydomonas reinhardtii.
TITLE Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
  Chlamydomonadaceae; Chlamydomonas.
JOURNAL 1 (bases 1 to 560)
MEDLINE Chlamydomonadaceae; Chlamydomonas.
COMMENT Chlamydomonas reinhardtii.
  The First Laboratory for Plant Gene Research
  Kazusa DNA Research Institute
  Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
  Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
  Location/Qualifiers
  1..560
  /organism="Chlamydomonas reinhardtii"
  /strain="C9"
  /db_xref="taxon:3055"
  /clone="ICL039h11_r"
  /note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
  XhoI; The cDNA library was constructed from cells cultured
  in a carbon stress acclimatized condition in which carbon
  dioxide concentration in the bubbling gas was changed from
  5% to 0.04%"
  92 a 186 c 173 g 109 t
BASE COUNT
ORIGIN
Query Match 17.1%; Score 272.8; DB 10; Length 560;
Best Local Similarity 99.3%; Pred. No. 4.9e-41;
Matches 274; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GCGCTGGACATCGTGATGGTTGCTGCTGAGGTGCGCCCTTGGTCCAAAGACGGCGGCCG 60
DB 285 GCGCTGGACATCGTGATGGTTGCTGCTGAGGTGCGCCCTTGGTCCAAAGACGGCGGCCG 344
QY 61 GCGATGTGACGTGGTGGCTGCTGCTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 120
DB 345 GCGATGTGACGTGGTGGCTGCTGCTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 404
QY 121 ATTGCCCTCGCTACGACAGTACGTGACGCTGCGGACACCTCGGTGCTGCTGCTGCTGCTGCTG 180
DB 405 ATTGCCCTCGCTACGACAGTACGTGACGCTGCGGACACCTCGGTGCTGCTGCTGCTGCTGCTG 464
QY 181 ATGGCGAGAGTCCGCTTACTTCCACTCCATCAAGAGGGGCTGCGACCGCGTGTGGATT 240
DB 465 ATGGCGAGAGTCCGCTTACTTCCACTCCATCAAGAGGGGCTGCGACCGCGTGTGGATT 524
QY 241 GACCACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 276
DB 525 GACCACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 560

RESULT 13
LOCUS AV631115
DEFINITION AV631115 Chlamydomonas reinhardtii 5% to 0.04% CO2 Chlamydomonas
  reinhardtii cDNA clone ICL089a12_r 5', mRNA sequence.
ACCESSION AV631115
VERSION AV631115.1 GI:10793749
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
  Chlamydomonadaceae; Chlamydomonas.
AUTHORS Chlamydomonas reinhardtii.
TITLE Chlamydomonas reinhardtii.
  The First Laboratory for Plant Gene Research
  Kazusa DNA Research Institute
  Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
  Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
  Location/Qualifiers
  1..560
  /organism="Chlamydomonas reinhardtii"
  /strain="C9"
  /db_xref="taxon:3055"
  /clone="ICL039h11_r"
  /note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
  XhoI; The cDNA library was constructed from cells cultured
  in a carbon stress acclimatized condition in which carbon
  dioxide concentration in the bubbling gas was changed from
  5% to 0.04%"
  92 a 186 c 173 g 109 t
BASE COUNT
ORIGIN
Query Match 17.1%; Score 272.8; DB 10; Length 560;
Best Local Similarity 99.3%; Pred. No. 4.9e-41;
Matches 274; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GCGCTGGACATCGTGATGGTTGCTGCTGAGGTGCGCCCTTGGTCCAAAGACGGCGGCCG 60
DB 285 GCGCTGGACATCGTGATGGTTGCTGCTGAGGTGCGCCCTTGGTCCAAAGACGGCGGCCG 344
QY 61 GCGATGTGACGTGGTGGCTGCTGCTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 120
DB 345 GCGATGTGACGTGGTGGCTGCTGCTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 404
QY 121 ATTGCCCTCGCTACGACAGTACGTGACGCTGCGGACACCTCGGTGCTGCTGCTGCTGCTGCTG 180
DB 405 ATTGCCCTCGCTACGACAGTACGTGACGCTGCGGACACCTCGGTGCTGCTGCTGCTGCTGCTG 464
QY 181 ATGGCGAGAGTCCGCTTACTTCCACTCCATCAAGAGGGGCTGCGACCGCGTGTGGATT 240
DB 465 ATGGCGAGAGTCCGCTTACTTCCACTCCATCAAGAGGGGCTGCGACCGCGTGTGGATT 524
QY 241 GACCACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 276
DB 525 GACCACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 560

TITLE Generation of expressed sequence tags from low-CO2 and high-CO2
  adapted cells of Chlamydomonas reinhardtii
JOURNAL DNA Res. 7 (5), 305-307 (2000)
MEDLINE 20539644
COMMENT Contact: Erika Asamizu
  The First Laboratory for Plant Gene Research
  Kazusa DNA Research Institute
  Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
  Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
  Location/Qualifiers
  1..393
  /organism="Chlamydomonas reinhardtii"
  /strain="C9"
  /db_xref="taxon:3055"
  /clone="ICL089a12_r"
  /note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
  XhoI; The cDNA library was constructed from cells cultured
  in a carbon stress acclimatized condition in which carbon
  dioxide concentration in the bubbling gas was changed from
  5% to 0.04%"
  64 a 145 c 137 g 47 t
BASE COUNT
ORIGIN
Query Match 16.8%; Score 267; DB 10; Length 393;
Best Local Similarity 100.0%; Pred. No. 5.5e-40;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1327 ACCGGCTTCCACATGGCGCCCTGAACCCGACACGCTGACGAGGCTGACGCCGACGCC 1386
DB 1 ACCGGCTTCCACATGGCGCCCTGAACCCGACACGCTGACGAGGCTGACGCCGACGCC 60
QY 1387 CTGGCGCCACCGTGGCGCGTGGCCAGCGAGGTGTTTTCGGGGCGCGCTACCCCGAGATG 1446
DB 61 CTGGCGCCACCGTGGCGCGTGGCCAGCGAGGTGTTTTCGGGGCGCGCTACCCCGAGATG 120
QY 1447 GTGGCCAACTGATCATCAGCCAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1506
DB 121 GTGGCCAACTGATCATCAGCCAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 1507 CTGCTGGAGGAGTGTGTACGCAAGGGCGGCTGGCCACCGCCCAAGAGAGGAGGATC 1566
DB 181 CTGCTGGAGGAGTGTGTACGCAAGGGCGGCTGGCCACCGCCCAAGAGAGGAGGATC 240
QY 1567 AAGTGCCCGCTTGGCCGAGAGATCCCC 1593
DB 241 AAGTGCCCGCTTGGCCGAGAGATCCCC 267

RESULT 14
LOCUS BE024926
DEFINITION BE024926 552 bp mRNA linear EST 06-JUN-2000
  Chlamydomonas reinhardtii CC-1690, normalized, Lambda zap II
  Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION BE024926
VERSION BE024926.1 GI:8287367
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
  Chlamydomonadaceae; Chlamydomonas.
AUTHORS Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,
  McDermott,J.P., Silflow,C., Stern,D. and Surzycki,R.
  Analyses of the Chlamydomonas reinhardtii Genome: A Model,
  Unicellular System for Analyzing Gene Function and Regulation in
  Vascular Plants: project phase 2
  Unpublished (2000)
  Contact: Elizabeth H. Harris
  DCMB Box 91000
  Duke University
  Durham, NC 27708-1000, USA
  Tel: 919 613 8164

```

Fax: 919 613 8177

Email: chlamy@duke.edu

FEATURES Location/Qualifiers

source

1..552

/organism="Chlamydomonas reinhardtii"

/strain="CC-1690 wild type mt+ 21gr"

/db_xref="taxon:3055"

/clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap

II"

/note="vector: pBluescript II SK-; Site_1: EcoRI; Site_2:

XhoI; This library, constructed by John Davies and Jeffrey

McDermott, combines cDNAs from CC-1690 cells grown to

mid-log phase in TAP (acetate-containing) medium in the

light, TAP medium in the dark, HS (minimal) medium in

ambient levels of CO2 and HS medium bubbled with 5% CO2.

PolyA mRNA was purified from each sample, pooled and cDNA

synthesized. The cDNA was directionally cloned into lambda

ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites.

pBluescript II SK- plasmids were excised from the lambda

ZAP clones by superinfection with ExAssist (Stratagene)

phage. The library was normalized using method 4 described

in Bonaldo et al (1996) Genome Research 6: 791-806."

BASE COUNT 91 a 185 c 167 g 109 t

ORIGIN

Query Match

Best Local Similarity 15.8%; Score 251.8; DB 10; Length 552;

Matches 253; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCGGTGGACATCGTATGTTGCTGCTGAGGTGCGCCCTTGCTCAAGACGGCGGCTG 60

Db 298 GCGGTGGACATCGTATGTTGCTGCTGAGGTGCGCCCTTGCTCAAGACGGCGGCTG 357

QY 61 GCGATGCTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120

Db 358 GCGATGCTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 417

QY 121 ATTGCCCCCTCGCTAGCAGTACGCTGACGCTGCGGACCTGCGTGGTGGTGGT 180

Db 418 ATTGCCCCCTCGCTAGCAGTACGCTGACGCTGCGGACCTGCGTGGTGGTGGT 477

QY 181 ATGGCGGAGAGGTCCGCTACTTCCATCCATCAAGAGGGCGTGCACCGGTGTGGATT 240

Db 478 ATGGCGGAGAGGTCCGCTACTTCCATCCATCAAGAGGGCGTGCACCGGTGTGGATT 537

QY 241 GACCAACCCCTGGTTC 255

Db 538 GACCAACCCCTGGTTC 552

RESULT 15

AV641583

LOCUS

AV641583 Chlamydomonas reinhardtii 517 bp mRNA linear EST 15-DEC-2000

cDNA clone HCL036g08_r 5', mRNA sequence.

ACCESSION AV641583

VERSION AV641583.1 GI:10784911

KEYWORDS EST.

SOURCE Chlamydomonas reinhardtii.

ORGANISM Chlamydomonas reinhardtii

Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

Chlamydomonadaceae; Chlamydomonas.

REFERENCE 1 (bases 1 to 517)

AUTHORS Asamizu E., Miura K., Kucho K., Inoue Y., Fukuzawa H., Ohya K.,

Nakamura Y. and Tabata S.

TITLE Generation of expressed sequence tags from low-CO2 and high-CO2

adapted cells of Chlamydomonas reinhardtii

JOURNAL DNA Res. 7 (5), 305-307 (2000)

MEDLINE 20539644

COMMENT

Contact: Erika Asamizu

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

FEATURES

source

1..517

/organism="Chlamydomonas reinhardtii"

/strain="C9"

/db_xref="taxon:3055"

/clone_lib="HCL036g08_r"

/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:

XhoI; The cDNA library was constructed from cells cultured

in a medium with bubbling air containing 5% carbon

dioxide"

BASE COUNT 82 a 176 c 159 g 100 t

ORIGIN

Query Match

Best Local Similarity 15.7%; Score 250.2; DB 10; Length 517;

Matches 252; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCGGTGGACATCGTATGTTGCTGCTGAGGTGCGCCCTTGCTCAAGACGGCGGCTG 60

Db 263 GCGGTGGACATCGTATGTTGCTGCTGAGGTGCGCCCTTGCTCAAGACGGCGGCTG 322

QY 61 GCGATGCTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120

Db 323 GCGATGCTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 382

QY 121 ATTGCCCCCTCGCTAGCAGTACGCTGACGCTGCGGACCTGCGTGGTGGTGGT 180

Db 383 ATTGCCCCCTCGCTAGCAGTACGCTGACGCTGCGGACCTGCGTGGTGGTGGT 442

QY 181 ATGGCGGAGAGGTCCGCTACTTCCATCCATCAAGAGGGCGTGCACCGGTGTGGATT 240

Db 443 ATGGCGGAGAGGTCCGCTACTTCCATCCATCAAGAGGGCGTGCACCGGTGTGGATT 502

QY 241 GACCAACCCCTGGTTC 255

Db 503 GACCAACCCCTGGTTC 517

Search completed: June 3, 2003, 15:27:33

Job time : 1560.22 secs

This Page Blank (uspto)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2003, 09:49:44 ; Search time 38.3196 Seconds
(without alignments)
12748.991 Million cell updates/sec

Title: US-09-980-771A-8

Perfect score: 1593

Sequence: 1 gcgcggacatcgtgatgt.....ccgttgccgagaagatcccc 1593

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents.NA.*

- 1: /cgn2.6/ptodata/1/ina/5A.COMB.seq.*
- 2: /cgn2.6/ptodata/1/ina/5B.COMB.seq.*
- 3: /cgn2.6/ptodata/1/ina/6A.COMB.seq.*
- 4: /cgn2.6/ptodata/1/ina/6B.COMB.seq.*
- 5: /cgn2.6/ptodata/1/ina/PCTUS.COMB.seq.*
- 6: /cgn2.6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	474.8	29.8	2267	4	US-08-679-645-25
2	447.4	28.1	2542	3	US-08-941-445A-6
3	276.2	17.3	2067	4	US-09-388-743-21
4	266.6	16.7	2274	4	US-09-388-743-17
5	259.6	16.3	2176	4	US-09-388-743-13
6	255.4	16.0	2202	4	US-09-388-743-1
7	218.2	13.7	2825	4	US-09-196-390-5
8	216.8	13.6	2097	3	US-08-941-445A-10
9	210.2	13.2	2380	1	US-08-572-951-3
10	207	13.0	2248	4	US-09-345-214-20
11	206.6	13.0	1798	4	US-09-345-214-16
12	206.6	13.0	2019	4	US-09-345-214-15
13	161	10.1	2007	3	US-08-941-445A-8
14	161	10.1	2085	1	US-08-572-951-2
15	118.2	7.4	2239	4	US-09-196-390-1
16	109.6	6.9	2348	4	US-09-388-743-5
17	108	6.8	1758	3	US-08-836-567-3
18	106.8	6.7	2418	4	US-09-388-743-25
19	106.6	6.7	2360	3	US-08-836-567-9
20	102.6	6.4	1528	4	US-09-345-214-6
21	102.6	6.4	1620	3	US-08-941-445A-20
22	102.6	6.4	1752	3	US-08-941-445A-12
23	102.6	6.4	2008	4	US-09-345-214-12
24	102.6	6.4	2491	4	US-09-345-214-5
25	101	6.3	2383	4	US-09-192-909-1
26	97.8	6.1	2990	1	US-08-572-951-1
27	97.4	6.1	4800	3	US-08-941-445A-4

Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 7, Appli
Sequence 1, Appli
Sequence 5, Appli
Sequence 11, Appli
Sequence 11, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 2, Appli
Sequence 5, Appli
Sequence 7, Appli
Sequence 1, Appli
Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-08-679-645-25
; Sequence 25, Application US/08679645
; Patent No. 6350934
; GENERAL INFORMATION:
; APPLICANT: Zwick, Michael G.
; APPLICANT: Edington, Brent E.
; APPLICANT: McSwiggen, James A.
; APPLICANT: Merlo, Patricia Ann Owens
; APPLICANT: Guo, Lining
; APPLICANT: Skokut, Thomas A.
; APPLICANT: Young, Scott A.
; APPLICANT: Folkerts, Otto
; APPLICANT: Merlo, Donald J.
; TITLE OF INVENTION: COMPOSITION AND METHODS FOR
; TITLE OF INVENTION: MODULATION OF GENE EXPRESSION
; TITLE OF INVENTION: IN PLANTS
; NUMBER OF SEQUENCES: 1263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: Los Angeles
; COUNTRY: California
; ZIP: U.S.A.
; 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/679,645
; FILING DATE: July 12, 1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,135
; FILING DATE: July 13, 1995
; APPLICATION NUMBER: 08/300,726
; FILING DATE: September 2, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 219/247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:
; LENGTH: 2267 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-679-645-25

Query Match 29.8%; Score 474.8; DB 4; Length 2267;
Best Local Similarity 59.9%; Pred. No. 6.7e-80;
Matches 978; Conservative 0; Mismatches 557; Indels 99; Gaps 7;
QY 5 TGGACATCGTGAATGTTCTGCTGAGTCCGCCCCCTTGTGCAAGACGGCGGCGCTGGGGG 64
DB 418 TGAACGTCGTCTCTGCGCGCGAGATGGCGCCGTGGAGCAAGACCGCGCGCTCGGCG 477
QY 65 ATGTGACTGGTGGCTGCTATTTAGCTGGTCAACGGCGGCGCACCGCTGCATGACCATG 124
DB 478 ACGTCTCGGGGGCTGCGCGCGCGCATGGCGCGAAATGGGACCGCTGTATGGTCTGT 537
QY 125 CCCCTCGCTAGACACAGTACGCTGACGGCTGGGACACCTCGGTGGTTCGTGGACAT---- 179
DB 538 CTCGCCGCTAGACACAGTACAGGACGCTGGGACACAGCGTCTGTCCGAGATCAAGA 597
QY 180 -----CATGGGCGAGAGTTCGCTACTTCCACTCCATCAAGAGGCGGTGCACCGG 232
DB 598 TGGGAGACAGGTACGAGACGCTCAGGTTCTCCACTGTCTACAAAGCGCGAGTGGACCG 657
QY 233 TGTGATGATGACACCCCTGCTTCTGGCAAGGTCTGGGCAAGACCGGCTCCAGCTGT 292
DB 658 TGTTCGTGTACACCCACTGTTCCTGGAGAGGTTGGGAAAGACCGGAGGAGATCT 717
QY 293 ACGGCGCCGCTTCGCGCGTGTACTACTGTGCAACCAAGCGCTTCGCCCTGTTGTGCA 352
DB 718 ACGGCGCTGACGCTGGAACGGACTACAGGACCAAGCTGCGGTTCAGCCTGCTATGCC 777
QY 353 ACGGCGCTATTAGGCTGCCCGCTGCTCCCTTCGCG----- 390
DB 778 AGGCAGCACTTGAAGCTCAGAGGATCTCTGAGCTCAACCAACCACTACTTCTCCGGAC 837
QY 391 --CCGCGCGGAGACTGCGCTTCCTGGGCAAGCAAGCTGCGACTCGCGCTTCGCGCTCC 448
DB 838 CATACGGGAGGAGCTGCTGCTGCTGCAACGACTGGCACCGGCTCTCTCGTGT 897
QY 449 TGTGAAGGACGAGTACAGCCCAAGGCGGCTTCAACAGGCAAGTGGTGTCTGGCTA 508
DB 898 ACCTAAGAGCACTACCACTGCCAGGCTACAGGAGCAAGAGCGGCTTCTTGTCA 957
QY 509 TCCACAACTCGCTTCCAGGCGCGCATGTGGGAGGAGCTTCAAGGACACGAAGCTGC 568
DB 958 TCCACAACTCTCTACAGGCGCGGTTTCGCTTCCGACTACCGGAGCTGAACCTCC 1017
QY 569 CCCAGCGCGCTTTCACAACTGCGCTTCTCGSACGGCTATGCCAAGGTTTACACTGAGG 628
DB 1018 CGGAGAGATTCAAGTCTTCCTTCGATTTCATCGACGCTACGAGAAG----- 1064
QY 629 CCACCCCATGAGGAGGAGGAGAGACCCCGCTGACGGGAAGACCTTACAGAAGATCA 698
DB 1065 -----CCCCGGAAG-----GCCGGAAGATCA 1086
QY 689 ACTGGCTGAAGGTTGCTATTCGCGCGCGAGCAAGCTGTGACTGTGTCGCCCACTACG 748
DB 1087 ACTGGATGAAGCGCGGATCTCTGAGCGCGACAGGCTCTTCCCTGTCAGCCCTACTACG 1146
QY 749 CGACCGAGATCGTCTCGATGCGCGCGGGGTGTGGAGCTGGACACCGCTATCCGCGCA 808
DB 1147 CCGAGGAGCTCATCTCCGCAATCGCCAGGGCTGCGAGCTGCAACAATCATCGCGCTCA 1206
QY 809 AGGCAATTGAGGCAATTGTAACGGCATGGACATTTGAGAGTGGAAACCCCAAGACGACA 868
DB 1207 CCGGCATACCGGCTGCTCAACGGATGAGCTGACGAGTGGGACCCCGACGAGGACA 1266
QY 869 AGTCTGCTGCGGCTTACGACAGCAAGCTCTACCGCGCAAGGCGCGCGCAAGG 928

RESULT 2

US-08-941-445A-6
; Sequence 6, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanning
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

DB 1267 AGTACATCCCGCTGAAGTACGAGCTGTGCAAGCGCGTGGAGGCCAAGGCGCTCAACAAGG 1326
QY 929 AGGCGCTGACGCGCGAGCTGGCGCTGCTGTGGACCCCAACCGCCCGCTGTTCGCGCTCA 988
DB 1327 AGGCGCTGACGCGCGAGCTGGCGCTCCCGTGGACCGGAACATCCCGCTGTGGCGCTTCA 1386
QY 989 TCGCGCGCTGAGGAGCAAGAGGTGTGGACATCATCTGCGCGCGCTGCGCGCTGCGCAAGATCC 1048
DB 1387 TCGCGAGGCTGAAGAGCAAGAGGACCGAGTCTATGGCGCGCGCTATCCCGAGCTCA 1446
QY 1049 TCGCCACCCCAAG--GTGCAATCGCATCTTGGGTACCGGCAAGGCGCGCTACGAGA 1105
DB 1447 TGGAGATGTGAGAGCACTGCAATCTGTCTGGCACGGCAAGAGTTCGAGC 1506
QY 1106 AGCTGTGAACCCCATCGGCACCAAGTACAAGGGCGCGGCCAAGGCGTGTGCAAGTCT 1165
DB 1507 GCATGCTATGAGCGCGAGGAGAGTTCACAGGCAAGGTGGCGCGCTGGTCAAGTTC 1566
QY 1166 CGGCGCGCTGGCGACATGCTCACCGCGCGCGCGCTCATCTGCTGTGCTGCGCTCGCGCT 1225
DB 1567 ACGCGCGCTGGCGCACCATCATGGCGCGCGCGCGCTGCTGCGCGTCAACGCGCGCT 1626
QY 1226 TCGAGCGCTGGCGCTGATCCAGCTGACGCGCATGCTACGCTACGCTGCGCGTGGTAG 1285
DB 1627 TCGAGCGCTGGCGCTGATCCAGCTGACGCGCATGCTACGCTACGCTGCGCGTGGTAG 1686
QY 1286 CTTCCACCGCGCGCTGTGCGACACCGCTCAAGAGGCGCTGCGCGTGTGCGCGTTCACATGGGCG 1345
DB 1687 CGTCCACCGCTGAGTCTGCGACACCATCATCGAAGGCAAGACCGGTTCCACATGGGCG 1746
QY 1346 CCGTGA-----ACCCGACAGCTGAGAGGCTACGCGGACGCGCTGCGCGCGCGCGCG 1399
DB 1747 GCTCAGCTGCGACTGCAACGCTGTGGAGCGCGCGCGCTCAAGAGGTTGGCGCGCGCGCT 1806
QY 1400 TCGCGCGTGGCGAGGCTGTGCGCGCGCGCTACCGCGAGTACCGCGAGTGTGGCGCGCGCT 1459
DB 1807 TCGAGCGCGCTCAAGTGTGCGCGCGCGCTACGAGGAGTGTGTGAGGAGTGTGAGGAGTGTCA 1866
QY 1460 TCGAGCGAGCTGTCTGTGCTCAAGCGCGCGCGCTGCGCGAGTGTGGAGGCGCTGCTGAGGAGG 1519
DB 1867 TGATCAGGATCTCTCTGAGGCGCGCTGCGCGAGGAGTGTGGAGGAGTGTGCTGAGCG 1926
QY 1520 TGTGTACGCAAGGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 1579
DB 1927 TCGGCGTGGCG---GGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1983
QY 1580 CCGAGAGATCCCC 1593
DB 1984 AGGAGAGCGTGGCG 1997


```

: APPLICATION NUMBER: US/08/941,445A
: FILING DATE: 30-SEP-1997
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/026,855
: FILING DATE: 30-SEP-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Winner, Ellen P
: REGISTRATION NUMBER: 28,547
: REFERENCE/DOCKET NUMBER: 89-97
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (303) 499-8080
: TELEFAX: (303) 499-8089
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2542 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: not relevant
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHETICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: Oryza sativa
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 453..2282
: US-08-941-445A-6

Query Match      28.1%;   Score 447.4;   DB 3;   Length 2542;
Best Local Similarity 58.6%;   Pred No. 8.1e-75;
Matches 930;   Conservative 0;   Mismatches 561;   Indels 96;   Gaps

QY  5  TGGACATCGTGATGGTTGCTGTAGGTGCGCCCTTGTTCCAAAGACGGCGCGCCTGGGCG  64
    || || || || || || || || || || || || || || || || || || || || ||
Db   700  TGAACGTCGTGTTTCGTGGCGCGGAGATGGCCCCCTGGAGCAAGACGCGCGCCTCGGTG  759

QY  65  ATGTGACTGTGTGGCTTGCCTATTGACTGTGTCGAAGCGGGGCCACCGCTCATGACCATG  124
    || || || || || || || || || || || || || || || || || || || || ||
Db   760  ACGTCCCTGGTGGCTTCCGCCCTGCCCATGGCTGCGAATGGCCACAGGTCATGGTGAATC  819

QY  125  CCGCTCGCTACGACCAGTAGCTAGCTGAGCGCTGGGACACCTCGGTGGTCTGTGCACATCATTG  184
    || || || || || || || || || || || || || || || || || || || || ||
Db   820  CTCTCGGTACGACACAGTACAGGAGCGCTTGGATACACAGCGTTGTGCTGAGATCAAGG  879

QY  185  -----CCGAGAAGTCCCGCTACTTCCACTTCCATCAAGAAGGGCGGTGCACCGCG  232
    || || || || || || || || || || || || || || || || || || || || ||
Db   880  TTGCAGACAGGTTACGAGAGSGTGAAGTCTTTCATTGCTACAAGGCTGAGTGCACCGTG  939

QY  233  TGTGGATTGACCAACCCCTGGTTCCTGGCCCAAGGTCTGGGGCAAGACCGGCTCCAAAGTGTT  292
    || || || || || || || || || || || || || || || || || || || || ||
Db   940  TGTTCATTCGACCATFCGTCTATTCCTGGAGAAGGTTGGGAAAGACCGGTCGAGAAGATCT  999

QY  293  ACGGCCCCCGCTCCGGCGCTGACTACTCGGCAACCAACAAGCGTTCGCCCTGTCTGCCA  352
    || || || || || || || || || || || || || || || || || || || || ||
Db   1000  ACGGACCTGACACTGAGTGTGATTACAAGACACACAGATGCGTTTCAGCCTTCTTTGCC  1059

QY  353  AGGCCCGCTATTGAGGCTGCCCGCGTCTGCCCTTCGGCCCG----- 393
    || || || || || || || || || || || || || || || || || || || || ||
Db   1060  AGGCAGCACTCGAGGGCTCTAGGATCCCTAAACCTCAACAACACCCATACTTCAAAGAA  1119

QY  394  ----GGCAGGAGCTCGCTTTCGTGGGCCAAGCACTGCGACTTCGCCCTCGCTGCCCGTCC  448
    || || || || || || || || || || || || || || || || || || || || ||
Db   1120  CTTATGTTGAGGATGTGTGTTTCGTTCTGCAAGCACTGCGACACTGGCCCCACTGGCGAGCT  1179

QY  449  TGTCTGAGGACGAGTACCAGCCCCAAGGCCAGTTTCAACAAGGCCAAGTCGGTGTCTGGCTA  508
    || || || || || || || || || || || || || || || || || || || || ||
Db   1180  ACCTGAAGAACAACTACCAGGCCAATGGCATCTACAGGAATGCAAAAGGTTCTCTTCTGCA  1239

QY  509  TCCACAACATCGCCTTCCAGGCGCGCATGTGGAGGAGGCTTTTCAAGACACAGCAAGCTGC  568
    || || || || || || || || || || || || || || || || || || || || ||
Db   1240  TCCACACATCTCTACAGAGGCCGTTTCGCTTCGAGATTTACCTTGAGCTGAACCTCT  1299

QY  569  CCGAGGCCGCGCTTTGACAAGCTGGCCTTCTCGGAGCGGCTATGCCAAGGTTTACACTGAGG  628

```

; APPLICANT: Singletary, George
; APPLICANT: Zhou, Lan
; TITLE OF INVENTION: No. 6423886el Starch Synthese Polynucleotides and Their
; FILE REFERENCE: 1144
; CURRENT APPLICATION NUMBER: US/09/388,743
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 2067
; TYPE: DNA
; ORGANISM: Tulipa fosteriana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (58)...(1857)
US-09-388-743-21

Query Match 17.3%; Score 276.2; DB 4; Length 2067;
Best Local Similarity 51.9%; Pred. No. 4.2e-43;
Matches 820; Conservative 0; Mismatches 663; Indels 96; Gaps 5;

```
QY 1 GGCCTGGACATCGTATGTTGCTGAGGTGCGCCCTTGGTCCAAAGAGCGCGCCCTG 60
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 271 GGGATGAACCTGGTGTCTCGGGACCGAGACGGGTCCGTACAGCAAGACCGCGGGCTC 330
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 GCGATGTCACTGGTGGCTGCTATTGAGCTGGTCAAGCGCGGCCACCGCGTCATGACC 120
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 331 GGGATGTCTAGAGGGTTACGGCCGCCCTGGCGGCGAGAGGGCATCGGTCATGGTT 390
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 ATTGCCCTTCGTACGACCACTAGCTGACGCTGGGACACCTTCGGTGTCTGTCGACATC 180
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 391 GTCACCTCCGCGGTACGATCAGTACAAGGATGATGGGACACAACTGTGCTTGATGC 450
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 ATGGGC-----GAGAGTCCGCTACTTCCACTCCATCAAGAGCGCTGCAC 228
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 451 AAGTCGGGGATAAATGAGAGGGTTCCTTCTCCACTCCATCAAGAGGGGGTGGAT 510
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 229 CGCGTGTGGATTGACCAACCCCTGGTTCCTGGCAAGGTCTGGGGCAAGACCGCTCCAAG 288
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 511 AGGTCCTTATGATCACCTTGGTTCCTGAGAAGTTTGGGGAAACCGGTGGGAA 570
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 289 CTGTACGGCCCGCTCCGCGGCTGACTACTCTGGACAACCAAGCGCTTTCGCCCTGTC 348
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 571 TTGTATGGTCTCTTACTGSACTGATGATGATAAAGCAGCTACGGTTTCACTCTCTG 630
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 349 TGCAGGCGCTATTGAGGCTGCCCGTCTGCCCTTCGGCCCC----- 393
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 631 TGTACGGCTGTCTTGAGGCTCCAGAGTCTTAATCTCAACACAGTAAATATTTTCT 690
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 394 -----GGCAGGACTCGCTTCCTGGCCCAAGACTGCGACTCCCGCCCTGGTGCC 444
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 691 GGACCATATGTTGAAGATGCTGTTTATTGGAACGATTTGGCAGCTGACCTCTTCCA 750
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 445 GTCTCTGTAAGACGAGTACAGCCCAAGGCCAGTTCACCAAGGCCAAGTGGTCTG 504
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 751 TGTACTTGAAGAGTGTATTAATCAGAAGGATTATGAGAGTGCCTAAGGTTGCTTT 810
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 505 GCTATCCACAACTCCCTTCCAGCGCGCATGTGGAGGAGCTTTCAAGGACACCAAG 564
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 811 TGCATTTCATTAATATGCAATACCAAGCAGATTCGCCCTTCTGATTTCTGCTTCTCAC 870
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 565 CTGCCCCAGCGCCCTTTGACAGCTGGCTTCTCGAGCGCTATGCCAAGGTTTACACT 624
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 871 CTTCCAGACACATTAATCTCGTTGATTCTTCGATGATATACAAACCT----- 924
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 625 GAGGCCACCCCATGGAGGAGGACGAGAGCCCGCTGACGGGAAGACCTACAGAAG 684
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 925 -----GTGAAGGTAGAAA 939
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 685 ATCAACTGGCTGAGGTGGCATATTCGCGCGGACAAAGCTGGTACTGTGTCGCCCAAC 744
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 940 ATAAATTGGATGAGGCTGGAATATTTGAAGCGGACACTTGTAACTGTGAGCCGCTAT 999
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 745 TAGCGACCGAGATCGTCCGATCGCGCGCGGTTGAGCTGGACACCGTCATCCG 804
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1000 TATGCTTAAGAGCTCTCTCTGGAAGATAGAGTGTGAGTGTGACACAGTTCGCG 1059
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 805 GCAAGGGCATTTGAGGCGATTGAAACGGCATGAGCATTTGAGAGTGGAAACCCAGACC 864
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1060 TTGAGGGCGTCAAGGAATTTGAATGGATGGATGAATGAATGTGGAATCCATTGACA 1119
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 865 GACAAGTTCTCTGTCTGGCCCTACGACCAAGACAGGCTCTACGCCGCAAGCGCGCC 924
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1120 GACAAATTTATCACTGCAATACGATGCAACAAATGTAACAGAGCAAAACCTGTTAAT 1179
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 925 AAGGAGCGCTGACGCGCGAGCTGGCGCTGCTGTGACCCACCGCCCGCTGTCGCC 984
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1180 AAGCAAGAATTACAAGCAGAGTTGGCTTGCCTGTAGATCCAGACATTCCTGTTATAGTT 1239
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 985 TTCAATCGCGCGCTGGAGGAGAGAGGTGTGACATCATCATCTCTGGCGCGCTGCCAAG 1044
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1240 TTTGTTGAAGGCTGGAGGAGCAGAAAGTTTCAGATATTCTAGCTGACGAATTCAGAA 1299
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1045 ATCTGCGCACCCCAAGGTGACATCGCCATCTGGTACCGGCAAGCGCGCTACGAG 1104
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1300 TTAATGG---ATGAGACACGTTTCAGATCATTAATCTTGGAACTGCAAGAAACACCTCGAA 1356
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1105 AAGCTGTGAACGCCCATCGGCACCAAGTACAAGGGCGCGGCCCAAGGCGCTGGTCAAGTTC 1164
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1357 AAGGAGCTTGAAGAAATAGAAGAACAAATTTCCAGACAAGATGAGACTTGTTCGAAATTC 1416
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1165 TCGCGCGCCCTGGCGGCACATGCTCACCGCGCGCGCGCTCATGCTGTGGTCCCTCGCG 1224
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1417 AATGTTCCGTTGGCTCATATGATGATGGCTGGAGGTGATTTTATAATATTCCTAGTAGA 1476
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1225 TTCGAGCCCTGCGCCTGATCCAGCTGACGCCATGCACTACGGTACCGCTGCCGTGGTA 1284
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1477 TTTGAGCGCTGTGGCTTATTCAGCTTGAAGCATGAATATGGATGCCAGCATATGT 1536
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1285 GCCTCCACCGCGCGCTGTCGACACCGTCAAGAGGGCGTCAACCGCTTCCACATGGC 1344
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1537 TCCACACCGCTGCTGTGTAGACACAATCAGGGAAGGCTTCACCGGATTTTCACATGGGT 1596
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1345 GCCTGA-----ACCCGACAAAGCTTGACGAGGCTCACCGCGACGCCCTGGCGCCACC 1398
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1597 GCCTTCACTTGTAGTGTGAACHTTTGATCCGCTGATGTGGCGGAATTTGTTAAACT 1656
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1399 GTGCGCGCTGCCAGAGGCTTTTCGCGGGCGCGTACCCCGAGATGTTGGCCAACTGC 1458
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1657 GTAAAGAGGCCCTTAAGTCTATGGAACCTCCAGCCTTCAGCGAAATGGTTCAGAACTGC 1716
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1459 ATCAGCCAGGACCTGCTCTGTCACCGCGCGCGCGCGCGAGTGGAGGCGCTGCTGGAGAG 1518
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1717 ATGCTCAAGATCACTCATGGAAGAACTTCGCAAAAAATGGAAGAGCTACTCTCTGGGA 1776
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1519 GTGCTGTACGCGAGGCG 1537
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1777 CTGGAAGTCCAGCGACGCG 1795
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 4
US-09-388-743-17
; Sequence 17, Application US/09388743
; Patent No. 6423886
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Lan
; TITLE OF INVENTION: No. 6423886el Starch Synthese Polynucleotides and Their
; TITLE OF INVENTION: Use in the Production of New Starches
; FILE REFERENCE: 1144
; CURRENT APPLICATION NUMBER: US/09/388,743
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17

```
; LENGTH: 2274
; TYPE: DNA
; ORGANISM: Typha latifolia
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (115)...(1956)
US-09-388-743-17

Query Match      16.7%; Score 266.6; DB 4; Length 2274;
Best Local Similarity 52.9%; Pred. No. 2.5e-41;
Matches 828; Conservative 0; Mismatches 639; Indels 98; Gaps 8;

QY 1 GCGCTGACATCGTATGTTGCTGCTGAGTCCGCCCTTGTGTCACCAAGCGGCGGCTG 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 373 GGCATGAACCTTATCTTTGAGCTGAGATGCTCCATGGACAGACAGTGGAGGCTT 432
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 GGGGATGTGACTGGTGGCTGCTATGAGCTGGTCAAGCGCGCCACCGCTCATGACC 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 433 GGTGATGTTCTTGGAGGACTCCACCGCATTTGGCGCAATGGACATCGATTATGGTT 492
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 121 ATTGCCCTCGCTACACCACTAGCTGAGCGCTGGACACCTCGGTGGTGGACAT- 179
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 493 ATAGCCCAAGTATGATCAATACATGGATGCTTTGGGATACAGATGCTCTTGTGATTG 552
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 180 -----CATGGCGAGAGGTCCGCTACTTCCACTTCCATCAAGAGGCGCTGCAC 228
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 553 AAGTTGGGATAGGTGTGAACCGTGGCTTCTTCACTGCTATAAAGAGAGTGTGAT 612
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 229 CGGCTGTGATGACACACCCCTGGTTCCTGGCCAAAGTCTGGGCAAGACCGGCTCCAAG 288
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 613 CGAGTTTGTGCTGATCACCTATGTTCTTGGCAAGTCTGGGGAAACCTGTTGGGAAG 672
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 289 CTGTAGGCGCCCGCTCCCGCGCTGACTACCTGGACACCAAGCGCTTGGCCCTGTTTC 348
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 673 ATTTATGGTCTTAACACTGGAACAGACTATCAGGCAATCAGTACGCTTCAAGCTTCTA 732
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 349 TGCAAGCGCGCTATTGAGCTGCCCGCTGCTGCCCTTCGGCCG----- 393
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 733 TGCCAGGACGATTTGGAAGCTCTTAGAATCTTAATCTCAACACAGTAGTCTTCTCT 792
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 394 -----GGCGAGACGTGCTCTTCGTGGCCACGACTGGCAGCTCCGCGCTTGGTCC 444
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 793 GGTCTCTATTGGGGAAGATGTTATCTCTCATTTGCAATGATGGCACACTTCTCTTCGCCA 852
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 445 GTCTGCTGAAGGACGAGTACAGCCCAAGGCGGCTTCCACCAAGCCCAAGTGGTGTG 504
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 853 TGTCTACTTAAGAGCATGTGACCATCCCGTGGCATTTACCAAGACGCCAAGGTGTGCTTC 912
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 505 GCTATCCACACATCGCCTTCCAGGCGCCGCTATGTTGGGAGGAGGCTTTTCAAGGACAGAA 564
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 913 TGCAATTCACAAATATATACCAAGCTCGATTTCTCCCTCAGACTTCGAATTTCTCAAT 972
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 565 CTGCCCCAGGCGCCTTTGACAAGCTGGCCTTCTCGGAGGCTATGCCAAGGTTTACACT 624
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 973 CTTCGCGAG-----AATTTCAATCT 993
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 625 GAGGCCACCCCATGAGGAGACAGAGCCCGCTCAGCGGGAAGACCTACAGAG 584
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 994 TCTTTTCAGTTTCATTGATGGGTACACACAGCC-----TGTGAAGGAATGAAG 1041
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 695 ATCAACTGGCTGAAGGGTGGCATTTATCGCGCGCGACAAAGCTGCTGACTGTGCGCCAC 744
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1042 ATAAATTTGATGAAGCAGGAATCTAGCAATCAGACAGGCTTTTACAGTGTAGTCCATAT 1101
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 745 TACGGCAGCGAGATCGTCCGATGCGCGCGGTGTGGAGCTGGACACCGCTCATCCGC 804
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1102 TATGCACAGAGCTCTTTTCAAGAGAAGAAAGGGAGTGCAGTTGGACAACATTTTGGCT 1161
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 805 GCAAGGCGATTTAGGGCATTTGTAACGCGCATGGACATGAGAGTGGAAACCCAGACC 864
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1162 GTGACCAAGTATCACAGGAATTTGTAATGGAATGGAATGTAATGAGTGAATCCATTAA 1221
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 865 GACAAGTTCCTGCTGCGCCCTTACGACCAAGACGCTTACGCGCGCAAGSCCGCGCC 924
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 1222 GACAAGTATATTTCTGTATTAATATGATGCAAAACCTTAATGGAAGCAAAAGCCTTTAA 1281
QY 925 AAGGAGGCCCTGCAGGCGG-AGCTGGGCGCTGCTGTGGACCCACCCGCGCCCTGTTCGC 983
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1282 AAGGAAGCATTTGCAAGCTGAAGTTGGCTTGCCTGTAGACAGGAGACATCCCTGTAATGT 1341
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 984 CTTATCGCGCGCTGGAGGAGCAAGAGGTGTGGACATCATCTGCGCGCGCTGCCCAA 1043
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1342 ATTCAAT-GGAAGACTAGAGGAGCAAGAGGATCAGATATTCTAGCAGCATCCATTCAGA 1400
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1044 GATCCTGGCCACCCCAAGCTGAGATCGGCATCTGGGTACCGGCAAGGCGCGCTACGA 1103
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1401 GATCATGG---ATGAGAAATGTTCAAGCTAATTAATTTCTTGGAACTGGCAAGAGAAATGGA 1457
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1104 GAAGCTGGTGAACGCCCATCGGCACCAAGTACAAAGGGCGCGCAAGGGCGTGGTCAAGTT 1163
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1458 GAATCAGCTTCAGAGTATGAGGAATGTTCCCGGACAAGGTGAGGCGAGTTATGAAGTT 1517
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1164 CTGCGCGCCCTGGGCGACAIGTTCACCGCGCGCGGCGACTTCAIGTGTGGTCCCTCGCG 1223
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1518 CAATGCTCCCTTAGCTCACCAGATGACGGCGGAGCTGATATAATTGTCATTCCAACTCG 1577
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1224 CTTGAGCCCTGCGGCTGATCCAGCTGCAGCCATGCATGCTACCTGACCTGCGCGTGGT 1283
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1578 ATTCGAACCAATCGGCGCTTATCCAGTTGCAAGGCATGCAATATGGAACGCGCTTCTGCGTG 1637
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1284 AGCCTCACCGCGGCTGTGTCGACACCGTCAAGGAGGCGTCAAGGCGTCCACATGGG 1343
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1638 TTCTCCACTGTGGGCTTGTGGACAGGTGAAGAGGCAAACTGGATTCCATATGGG 1697
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1344 CGCCCTGA-----ACCCGACAAGCTGAGGCTGACGCGGACGCGCTTGGCGCGCAC 1397
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1698 ACCTTTCACTGTGTAATGCGAAGTGTGCTGATCTTCTGATGTAAGAAAGTTGTCACAAC 1757
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1398 CGTGGCGCTGCAGGAGGTGTTTGGGCGCGCGCTACCCCGAGATGTTGGCCAACTG 1457
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1758 TGTGAACCGGCACTTAAGTCTATCGGCACACCTGCTTTGAGGATATGATCAAGAAATG 1817
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1458 CATCAGCAGGAGCTGCTCTGTTGTCACAGCCCGCCAGAGTGGAGGCGCTGCTGGAGGA 1517
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1818 CATGGCACAGAGCTCTCATGGAGGCGCTGCAAGAACTGGAGCAAGTTCTACTGAA 1877
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1518 GGTGG 1522
   |||
Db 1878 CTTGG 1882
```

RESULT 5

```
US-09-388-743-13
; Sequence 13, Application US/09388743
; Patent No. 6423886
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Ian
; TITLE OF INVENTION: No. 6423886el Starch Synthase Polynucleotides and Their
; FILE REFERENCE: 1144
; CURRENT APPLICATION NUMBER: US/09/388,743
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 2176
; TYPE: DNA
; ORGANISM: Canna edulis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)...(1928)
US-09-388-743-13

Query Match      16.3%; Score 259.6; DB 4; Length 2176;
Best Local Similarity 51.6%; Pred. No. 5e-40;
```

```
Matches 795; Conservative 0; Mismatches 649; Indels 96; Gaps 5;
QY 5 TGGACATCGTATGCTGCTGAGTCCGCCCTTGGTCCAAAGACGGCGCCCTGGGGG 64
Db 349 TGAACCTGGTGTGTTGTTGTTGAGGTAGTCCCAAGAGCAAACTGGGGCCCTGGGG 408
QY 65 ATGTGACTGGTGGCTGCTATGTAGCTGGTCAAGCGGGCCACCGGCTCATGACCATG 124
Db 409 ATGTGCTTAGAGGATTCGCAACCTGCTATGGCTGCAATGGGCACAGGGTCATGACGGTGG 468
QY 125 CCCCTGCTACACAGCTAGCTGAGCGCTGGGACACCTCGGTGGTGGTGGACAT----- 179
Db 469 TGGCAGCATATGACCAATAAAGATATCTGGGATACAACTGTCAGTTGAGTTAAAG 528
QY 180 -----CATGGCGAGAGGTCCGCTACTTCCACTCCATCAAGAGGGCGTGCACCGG 232
Db 529 TTGGGATAAGATTGAACTGTCGCTCTTCCACTGCTACAAAGGGAGTTGATCGGG 588
QY 233 TGTGATTACACACCCCTGGTTCCTGGCCAAAGTCTGGGGCAAGACCGGCTCCAGCTGT 292
Db 589 TTTTGTGATCACCCCTATGTTCTCGAGAAGGTTTGGGGAAACAGAGGAAATAT 648
QY 293 ACGGCGCCCTCGCGCTGACTACCTGGACAACACAGCGCTTCGCGCTGTTCTGCA 352
Db 649 ATGGTCCCTTTACGAACAGATTATGACAGCAATCACTAAGATTCAGCCCTTTGTGCC 708
QY 353 AGGCGCTATTGAGGTGCGCGCTGCTGCTCCCTTCGCGCCGCGGAGGACTGC----- 405
Db 709 TGGCAGCTCTGGAAGCTCCAGACTTCTAAATCTCAACACAGCAATACTATTCTGGAC 768
QY 406 -----GTCCTGGGCAAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 828
Db 769 CATATGGAGATGATGTTGTTTATTTGCAAGATTTGCAATTTGCTACTGCTGCTGCT 828
QY 449 TGTGAAGACGAGTACACGCCCAAGCGCCAGTTCACCAAGGCCAAGTGGTCTGGCTA 508
Db 829 ACTTGAACATATGATACCAATACATATGATTTACATGATGATTAAGTTGCAATTTGCA 888
QY 509 TCCACAACATCGCTTCCAGGCGCGCATGTGGGAGGAGGCTTTTCAAGGACACCAAGCTGC 568
Db 889 TTTCAATATATGCTTACAGGCGCGATTTGCTTTTCGGACTTTGAACCTTTAATCTCC 948
QY 569 CCGAGCGCCCTTTGACAGCTGGCGCTTCTCGGACGGCTATGCCAAGGTTTACACTGAGG 628
Db 949 CCAATAAATTTAAATCTTCAATTTGATTTGATGGATGGATGATGACAAACCT----- 998
QY 629 CCACCCCATGGAGGAGGAGAGACGCCCGCTGACGGGAAAGACCTTACAAAGATCA 688
Db 999 -----GTGAAAGAGGAAATTA 1017
QY 689 ACTGCTGAAGGTGTCATTTATCGCGCGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 748
Db 1018 ATTGGATGAAGCTGGAATAATAGATGTGATAGGTGCTTGACCGTGGAGCCCATATTATG 1077
QY 749 CGACCGATCGCTGCGGATGCGCGCGGGTGGAGCTGGACACCGTTCATCCGCGCCA 808
Db 1078 CCAAGAGCTTGTCTCAGGGGTAGAGAAGGTTGTTGAGTTGGCAATATCTCGCGCATGA 1137
QY 809 AGGCAATTGAGGCAATGTGAACGGCATGGACATTTGAGGAGTGGAAACCCCAAGACCGACA 868
Db 1138 AAACCATCTGTGGAATAGTAATAGGATGGACACACGAGGAGTGGATCCATTAACAGACA 1197
QY 869 AGTTCCTGCTGCGCCCTACGACACAGACGCTCTACGCGGCAAGCGCGCGCGCAAGG 928
Db 1198 AATATATTTCAAACTACGATGCAACAACCTGATTTGGATGCAAAACCTCTCTGTAAGG 1257
QY 929 AGCCCTTGAGGCGAGCTGGGCTGCTGCTGGACCCCGCGCCCGCTTTCGCTTCA 988
Db 1258 AAGCTTTGCAAGTGTGTTGGGCTGCTGTTAAACAAAACAAAGCTTGTGTTGGCGCTTG 1317
QY 989 TCGGCGCCCTGGAGGACGAAGGTTGGACATCATCTTGGCGCCCTGCGCCCAAGATCC 1048
Db 1318 TTGAAGACATAGATGACCAAGGCTCAGACATTTAGCTGACGAATTCAGACATTC 1377
```

RESULT 6

US-09-388-743-1

; Sequence 1, Application US/09388743

; Patent No. 6423886

; GENERAL INFORMATION:

; APPLICANT: Singletary, George

; APPLICANT: Zhou, Lan

; TITLE OF INVENTION: No. 6423886el Starch Synthese polynucleotides and Their

; TITLE OF INVENTION: Use in the Production of New Starches

; FILE REFERENCE: 1144

; CURRENT APPLICATION NUMBER: US/09/388,743

; CURRENT FILING DATE: 1999-09-02

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1

; LENGTH: 2202

; TYPE: DNA

; ORGANISM: Curcuma zedoaria

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (130)...(1974)

US-09-388-743-1

Query Match

Best Local Similarity 16.0%; Score 255.4; DB 4; Length 2202;

Matches 799; Conservative 0; Mismatches 646; Indels 96; Gaps 6;

```
QY 5 TGGACATCGTATGCTGCTGAGTTCGCCCTTGGTCCAAAGACGGCGCCCTGGGGG 64
Db 395 TGAACCTGATCTTGTGTTGTTGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 454
QY 65 ATGTGACTGGTGGCTGCTTATGAGCTGGTCAAGCGGGCCACCGGCTCATGACCATG 124
Db 455 ATGTCTTGGAGGTTTCCACCGCCCATGGCGGCAAGGACACAGGTTGATGACTATAG 514
QY 125 CCCCTGCTACGACACGACTAGCTGACCGCTGGGACACCTCGGTGGTGGTGGTGGTGG 184
Db 515 CACGCGGACATGACCAATAACAAGATGGATGGGATACGGGCTGCTTTTGTGAGTTGAAAG 574
```

QY 185 -----GCAGAGGTCCTACTTCCATCCATCAGAGGGCGTCCACCGG 232
DB 575 TTGGTATAGAAATGTCCTTTTCCACTGCTACAAAGGGGAGTTGATCGGG 634
QY 233 TGTGATGACCAACCCCTGTTCTGCGCAAGTCTGGGCAAGCGGCTCCAACTGT 292
DB 635 TGTGTTGGATCACCCTCTCTTCTGAGAGGTTTGGGAAAACCTGAGGAAATAT 694
QY 293 AGGCCCCCGCTCGCGCTGCTACTGAGCAACCAAGCGCTTCGCCCTTCTGCA 352
DB 695 ATGGTCTCTCAGAACTGATATGAGAACAAACAGCTAAGTTCTGTTCTGTT 754
QY 353 AGGCGCTATTGAGGTCGCCCGTCTG-----CCCTTCGGCC 391
DB 755 TGGCACTCTGAACTCCCAAGGTTCTGAATCCCAACAAATACAAATATCAATTCGGAC 814
QY 392 CGGGGAGGACTGCGTCTTCTGGCCCAAGCTGGGCACTCCGCCCTGTTGCCGCTCTGC 451
DB 815 CAAAAGGTGAAGATTATTCATTGCTAAGATGGCATCTGCTCTATTACCTTGTAAT 874
QY 452 TGAAGAC---CAGTACAGCCCAAGGGCCAGTTACCAAGGCAAGTGGTCTGGCTA 508
DB 875 TAAAGACCAATTATATCAAGCCCAATGGAATATACAAAATGCTAAAGTTGCTTTCTGCA 934
QY 509 TCCACAACATCCCTTCCAGGCGCATGTGGGAGGAGCTTTCAGAGGACACGAAGCTGC 568
DB 935 TTCATAATTCGTTATCAGGACGGTTTGCCTTTGAAGATTTCGCGTCTCAATCTCC 994
QY 569 CCCAGGCGCCCTTTCAGAGCTGGCTTCTCGACGGCTATGCCAAGGTTTACACTGAGG 628
DB 995 CTGATACATTCAAGTCTCTTTGATTTCATCGATGGCTATGCAAA-----1040
QY 629 CCACCCCATGAGGAGGACGAGAGCCCCCGCTGACGGGAAGACCTACAGAAATCA 688
DB 1041 -----ACCAATAAAGGAAGGAATCA 1063
QY 689 ACTGGTGAAGGTGGCAATTATCGCGCGCAAGCTGTGCTGCGCCCACTAGC 748
DB 1064 ACTGATGAGGCGGAATTATAGATCAGATCGTGCAATTGACTGTGACCCCATATATG 1123
QY 749 GCACCGAGATCGCTCGCGATCGCGCGGTGTGAGCTGGACACCGTCTATCCGCGCA 808
DB 1124 CCCAGGAATCGTCTCAGGAATCGATAGGGCGTGGAGTATGACATATACTCCGCTGA 1183
QY 809 AGGGCATTGAGGCGATTGTGACGCGATGACATGAGAGTGAACCCCAAGACGACA 868
DB 1184 AAACCATCTGTGGCATCAATAAATGAATGGACACCAACGAGTGGATCCCTCAACAGACA 1243
QY 869 AGTTCTCTGTCGCGCTACGACCAAGACAGCGTCTACCGCGCAAGCGCGCGCAAGG 928
DB 1244 ATATCATACAGCAATTACGACGCAACCACTGTATGAGGCAAGCCACTCAACAGG 1303
QY 929 AGGCGCTGAGGCGAGTGGCGCTGCTGTGGACCCCAACCGCCCGCTTCGCTTCA 988
DB 1304 AAGCTTTGCAAGCTGAGGTGGACTCGCGTCAACAGTAATCCCTGTGATAGTTTCA 1363
QY 989 TCGGCGCTGAGGACAGAGGTTGAGCATCATCTCGCGCCCTGCGCCCAAGATCC 1048
DB 1364 TTGGGACTAGAAAGAAAAGGTTTCAGACATTTCTAGCTGAAGCAATTCAAAAGTTCT 1423
QY 1049 TGGCCACCCCAAGTGCAGATCGCATCTCTGGTACCGCAAGCGCGCTACGAGAAGC 1108
DB 1424 T---CGATCAGATGTTCAAGTATAGTCTCGGTACTGTTAAAAGAGTTAGAGCGCC 1480
QY 1109 TGGTGAAGCCATCGGCACCAAGTACAAAGGCGCGGCAAGGCGTGTGTCAGTTCTCGG 1168
DB 1481 AACTTGTGATTCGAGGACGAGTTCCAGACAAATTCAGACCTCATATGAAGTTCAATA 1540
QY 1169 CGCCCTTGGCGCATGCTCACCGCGCGCGGCGGCTATGCTGCTGCTGCGCTTGG 1228
DB 1541 TTCTTTGGCTCATGGAATATGCGGGGTGCTGATATCTCTTTATTCAGCTAGGTTGG 1600
QY 1229 AGCCCTTGGCGCTGATCCAGCTGCGCATCTACCTACCGTACCGTCCCGTGGTAGCCT 1288

RESULT 7

US-09-196-390-5

; Sequence 5, Application US/09196390

; Patent No. 6307125

; GENERAL INFORMATION:

; APPLICANT: Block, Martina

; APPLICANT: Lorz, Horst

; APPLICANT: Lutticke, Stephanie

; APPLICANT: Walter, Lennart

; APPLICANT: Froberg, Claus

; APPLICANT: Kossmann, Jens

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES

; TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave

; STREET: 1251 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: United States of America

; ZIP: 10020

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; FILING DATE: 05/09/1996,390

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: DE 196 21 588.9

; FILING DATE: 29-MAY-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: DE 196 36 917.7

; FILING DATE: 11-SEP-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/EP97/02793

; FILING DATE: 28-MAY-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Haley, Jr., James F.

; REGISTRATION NUMBER: 27,794

; REFERENCE/DOCKET NUMBER: AGREVO-9

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 596-9000

; TELEFAX: (212) 596-9090

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2825 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

Db	1828	-----AGACCCGCGCATCGTCAACGGCATCGACAACATGAGTGGAAACCCCGAGG	1878
Qy	833	GCATGAGCATTTAGGAGTGGACCCCAAGACCGACAGTTCCTGTCTCGCCCTACGACC	892
Db	1879	TGGACGCCACCTCAAGTCCGGACGGCTACCAACATTCCTCTAGGACGC-----	1929
Qy	893	AGAACAGGCTTACGCCGCAAGGGCGCGCCCAAGAGGGCCCTGCAGCGCAGCTGGGCC	952
Db	1930	-----TGGACTCCGCAAGCGCACTGCAAGAGGCCCTGCAGCGCAGCTGGGCC	1980
Qy	953	TGCTTTGGACCCACCGCCCCCTTTTCGCTTCATCGCGCGCTTGGAGGACGAGAAGG	1012
Db	1981	TGCAGGTCCGCGCCGACGTGCGCTGCTCGGCTTCATCGCGCGCTGCAGCGGACAGAAG	2040
Qy	1013	GTGTGACATCATCTTGGCGCGCTTCCCAAGATCCTGGCCACCCCAAGGTGCAGATCG	1072
Db	2041	CGGTGGAGATCATCGCGGACGCTGCCCTGGATCGTG---AGCCAGAGCTGCAGCTGG	2097
Qy	1073	CCATCTTGGGTACCGGCAAGGGCGGCTACGAGAACTGTTGAACGCCATCGGCACCAAGT	1132
Db	2098	TGATGCTGGGCACCGGCGCCAGACCTGGAGAGCATGTCAGCACATTCGAGCGGGAGC	2157
Qy	1133	ACAGGGCCGCGCAAGGGCGTGGTCAAGTTCTCGGCGCCCTTGCSCACATGCTCACCG	1192
Db	2158	ACCACGAAAGTTCGCGGGTGGTGGGGTTCTCGCTGGCGCTTGGCGCACCGGATPCAGG	2217
Qy	1193	CGGCGCGCACTTCATGCTGGTGGCCCTCGCGCTTCAGCCCTTGGCGGCCTGATCCAGCTGC	1252
Db	2218	CGGGGGGACGGCGTCTCATGCCCTCCCGGTTTCGAGCGTTCGCGGCTGAACCACTCT	2277
Qy	1253	ACGCCATPGACTACGTTACGTTCGCGGTGAGCTTCACACGGGCGGCTGTTGCACACCG	1312
Db	2278	ACGCCATGGCTTACGGCACCGTCCCGCTGCTGCAGCGCTTCGCGCGCCTCAGGACACCG	2337
Qy	1313	TCAGGAGGGCGTCAACGGCTTCACATGGGCGCTTGAACCCCGACAAAGTTCGACGAGG	1372
Db	2338	TGCGCGGTTTCGACCCCTTCAACCACTTCGGGCTTGGGTGGACGTTTCGACCGCGCCGAGG	2397
Qy	1373	CTGACGCCGACGGCCTTGGCGGCAACCGTTCGCGCGTTCACAGCAGGTGTTTCGGGCGCGCC	1432
Db	2398	CGCAACAGCTGATCGAGGGCGCTCGGGCACTGCTCCGCACTTACCGAGACTTCAAGGAGA	2457
Qy	1433	GCTACCCCGAGATGGTGCCCACTGCATCAGCCAGGACCTGTCCTGTTCCAGGCGCGCC	1492
Db	2458	GCTGGAGGGCCCTCCAGGAGCGGGCATGTCGACGAGACTTCAGCTGGGAGCACGCGCCCA	2517
Qy	1493	AGAGTGGGAGGGCGCTGCTGGAGGAGTGGTGTAC	1527
Db	2518	AGCTCTACGAGGACGCTCTCGTCAAGGCCCAAGTAC	2552

```

RESULT 8
US-08-941-445A-10
; Sequence 10, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; City: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```


Db 2048 ARWSNTGGMGNCNTGYMGNMGNMGNATGCGNARGAYYTNHSNTGGGAYCAYGCMG 2107
QY 1490 CCCACAAGTGGAGGCGCTGCTGGAGGAGGTGGTGTAC--GGCAAGGGCGGCTGGCCAC 1547
Db 2108 CNGTNTATGARGAYGTYNTGNTNAARGCNAARTAYCARTGGTTRGCGNAAYTRYING 2167
QY 1548 CGCCAAAGAGGAGATCAAGGTCCGCTGCC 1581
Db 2168 CNACNMGNMGNMGNWSNTGYMGNMGNACNTGGAC 2201

RESULT 10
US-09-345-214-20
; Sequence 20, Application US/09345214
; Patent No. 6392120
; GENERAL INFORMATION:
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Broglie, Karen E.
; TITLE OF INVENTION: MODIFICATION OF STARCH BIOSYNTHETIC ENZYME GENE
; TITLE OF INVENTION: EXPRESSION TO PRODUCE STARCHES IN GRAIN CROPS
; FILE REFERENCE: BB-1147
; CURRENT APPLICATION NUMBER: US/09/345,214
; CURRENT FILING DATE: 1999-06-30
; EARLIER APPLICATION NUMBER: 060/094,436
; EARLIER FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 2248
; TYPE: DNA
; ORGANISM: Zea mays
US-09-345-214-20

Query Match 13.0%; Score 207; DB 4; Length 2248;
Best Local Similarity 55.1%; Pred. No. 2.8e-30;
Matches 477; Conservative 0; Mismatches 370; Indels 18; Gaps 3;

QY 678 CAAGAAGATCAACTGGCTGAAGGGTGGCATTTATCGCCCGCACACAGCTGGTGTGTC 737
Db 1331 CGACACGCCCAACATCTTTGCGCGGGTCTGAAGATGGCAGACCGGGTGTGACTGTGTC 1390

QY 738 GCCCAACTACCGACCGACGATCGCTGCGGATGCGCGCGGTGTGGAGCTGGACACCGT 797
Db 1391 CCGCGGTACCTGTGGAGCTGAACAGTGAAGCGGCTGGGCGCTCCAGACATCAT 1450

QY 798 C---ATCGCGCCCAAGGCATTTAGGGGATTTGAAGCGCATGGACATTTAGGAGTGGAA 854
Db 1451 CCGTTCTAACGACTGGAAGATCAATGGCATCTGTAACGGCATCGACACCAGGAGTGGAA 1510

QY 855 CCCCAAGACGACAAAGTTCTCTGTGCGCCCTACGACACGACAGCG-----T 902
Db 1511 CCCCAAGTGAAGTGCACCTGCGGTGCGGCGGTACACCACTACTCCCTCGAGACACT 1570

QY 903 CTACCGCGCAAGCCCGCCGCAAGGAGCGCTTCAGCGCCGAGCTGGGCGCTGCTGTGGA 962
Db 1571 CGACGCTGGAAGCGGAGTGAAGCGGCGCTTCAGCGGAGCTGGGCGCTGGAGTGGC 1630

QY 963 CCCACCGCCCGCTGTGCGCTTCATCGCGCGCTTCAGGAGGAGGAGGTGTGGACAT 1022
Db 1631 CGACGAGCTGCGCTGTGCGCTTCATCGCGCGCTTCAGGAGGAGGAGGTGTGGACAT 1690

QY 1023 CATCTTGGCGCCCTGCGGAGATCTTGGCCACCCCAAGTGCAGATCGGCATCTCTGGG 1082
Db 1691 CATCGGAGACCGATGCGGTGATC---GCGGGGCGAGCGTGCAGCTGTGATGCTGGG 1747

QY 1083 TACCGCAAGCCCGCTTACGAGAAGCTGGTGAACCGCATCGGACCAAGTACAGAGCGCG 1142
Db 1748 CACGCGGCGCGCACTTGGACCAAGTGTGACACATTTGAGCGGGAGCATCCCAACAA 1807

QY 1143 CGCCAAGGGGTGTCAAGTTCTCGGCGCCCTTGGCGCACATGCTACCGCGCGCGCGCA 1202
Db 1808 GGTGCGGGGTGGGTGCGGTTCCTGGTTCCTGCTTATGGCGCATCGCATCGGCGCGCGCA 1867

QY 1203 CTTCAAGTGGTGGCCCTCGCGCTTCGAGCCCTCGGCGCTGATCCAGCTGCACGCCATGCA 1262
Db 1868 CGTGTGTGTGATGCCCTCGCGCTTCGAGCCCTCGGCGCTGAACCACTCTACGCAIGGC 1927

QY 1263 CTACGCTACCGTGGCGGTGTAGCTTCACCGCGCGCTGTGGACACCGTCAAGGAGG 1322
Db 1928 ATACGCGACCGCTCGCTGTGTGTCACGCGCTGGCGGCTCAGGGACACCGTGGCGCGTT 1987

QY 1323 CGTCACCGGCTTCCACATGGCGCGCTGAACCCGCAAGCTGGAGAGGCTCACGCCGA 1382
Db 1988 CGACCGCTTCAGCGAGCGCGGCTCGGTGGACTTTTGACCGCGCGGAGCCCAAGCT 2047

QY 1383 CGCCCTGGCGCCACCGTTCGCGCTGCCAGCGAGTGTGTTGGGGCGCGCTACCCCGA 1442
Db 2048 GATCAGGCGCTCAGGCACTGCTCGACACGTACCGGAACCTACGAGGAGCTGGAAGAG 2107

QY 1443 GATGTTGGCAACTGCATCAGCAGGACCTGCTCTGTGTCACAGCCCGCCAGAGTGGGA 1502
Db 2108 TCTCAGGCGCGCGCATGTCGAGGACCTCAGCTGGGACCGCGGCTGAGCTCTACGA 2167

QY 1503 GGGCCTGCTGGAGAGGTGGTGTAC 1527
Db 2168 GGACGCTCTGTCAAGGCCAAGTAC 2192

RESULT 11

US-09-345-214-16/c
; Sequence 16, Application US/09345214
; Patent No. 6392120
; GENERAL INFORMATION:
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Broglie, Karen E.
; TITLE OF INVENTION: MODIFICATION OF STARCH BIOSYNTHETIC ENZYME GENE
; TITLE OF INVENTION: EXPRESSION TO PRODUCE STARCHES IN GRAIN CROPS
; FILE REFERENCE: BB-1147
; CURRENT APPLICATION NUMBER: US/09/345,214
; CURRENT FILING DATE: 1999-06-30
; EARLIER APPLICATION NUMBER: 060/094,436
; EARLIER FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 1798
; TYPE: DNA
; ORGANISM: Zea mays
US-09-345-214-16

Query Match 13.0%; Score 206.6; DB 4; Length 1798;
Best Local Similarity 55.0%; Pred. No. 3.3e-30;
Matches 476; Conservative 1; Mismatches 370; Indels 18; Gaps 3;

QY 678 CAAGAAGATCAACTGGCTGAAGGGTGGCATTTATCGCCCGCACACAGCTGGTGTGTC 737
Db 918 CGACACGCCCAACATCTTTGCGCGGGTCTGAAGTGGCAGACCGGTGGTGTGACTGTGAG 859

QY 738 GCCCAACTACCGACCGAGATCGCTGCGGATGCGCGCGGTGTGGAGCTGGACACCGT 797
Db 858 CCGCGGCTACCTGTGGAGCTGAAGACAGTGAAGCGGCTGGGCGCTCCAGACATCAT 799

QY 798 C---ATCGCGCCCAAGGCATTTAGGAGGATTTGTAACCGCATGGACATTTAGGAGTGGAA 854
Db 798 CCGTTCTAACGACTGGAAGATCAATGGCATCTGTAACGGCATTCGACACCAGGAGTGGAA 739

QY 855 CCCCAAGACCGCAAGTTCTCTGTGCGCCCTACGACCAAGCAACAGCG-----T 902
Db 738 CCCCAAGTGCAGTGTGCACCTGCGGTGCGGCTCGGACCACTACTCCCTCGAGACACT 679

QY 903 CTACCGCGCAAGCGCGCGCGAGGCGCTTCAGCGCGAGCTGGGCGCTGCGCTGTGGA 962
Db 678 CGACGCTGGAAGCGGAGTGCAGGCGCGCTTCAGCGGAGCTGGGCGCTGGAGTGGC 619

QY 963 CCCACCGCGCCCTGTGTGCGCTTCATGCGCGCTTCAGGAGGAGGAGGTGTGGACAT 1022

Db 618 CGACGAGTGGCGCTGCTGCGCTTTCATCGGGCTGTGATGACAGAAAGGGCGTGGACAT 559
QY 1023 CATCTGGCCGCTGCTCCAGATCTTGGCCACCCCAAGGTGAGATCGCATCTCTGG 1082
Db 558 CATGGGAGCGATGCGGTGGATC--GCGGGGAGGACGTGAGCTGGTGGTCTGG 502
QY 1083 TACGGGAAGCGCTACGAGAAGTGGTGAAGGCCATCGGCACCAAGTACAAAGGGCG 1142
Db 501 CACGGGGCGCGACCTGGAGCAATGCTCGACACTTGGAGCGGAGCATCCCAACA 442
QY 1143 CGCAAGGGCTGGTCAAGTCTTGGCGCCCTTGGCGCACATGTCACCGCGGGCGCGCA 1202
Db 441 GGTGGCGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 382
QY 1203 CTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1262
Db 381 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 322
QY 1263 CTACGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1322
Db 321 ATACGGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 262
QY 1323 CTTACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1382
Db 261 CGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 202
QY 1383 CGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1442
Db 201 GATCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 142
QY 1443 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1502
Db 141 TCTCAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 82
QY 1503 GGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1527
Db 81 GGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 57

RESULT 12

US-09-345-214-15
; Sequence 15, Application US/09345214
; Patent No. 6392120
; GENERAL INFORMATION:
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Broglie, Karen E.
; TITLE OF INVENTION: MODIFICATION OF STARCH BIOSYNTHETIC ENZYME GENE
; TITLE OF INVENTION: EXPRESSION TO PRODUCE STARCHES IN GRAIN CROPS
; FILE REFERENCE: BB-1147
; CURRENT APPLICATION NUMBER: US/09/345,214
; CURRENT FILING DATE: 1999-06-30
; EARLIER APPLICATION NUMBER: 060/094,436
; EARLIER FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 2019
; TYPE: DNA
; ORGANISM: Zea mays
US-09-345-214-15

Query Match 13.0%; Score 206.6; DB 4; Length 2019;
Best Local Similarity 55.0%; Pred. No. 3.3e-30;
Matches 476; Conservative 1; Mismatches 370; Indels 18; Gaps 3;

QY 678 CAAGAGATCAACTGGCTGAAGGTGGCAATATTCGCCCGCCGACAAAGCTGTGACTGTGTC 737
Db 887 CGAGACGCGCAACATCTTTGCGCGGGCTCTGAAGATGGCAGACCGGGTGTGACTCTAG 946
QY 738 GCCCAACTACGGGACGAGATGCTGCCGATGCCCGCGGGTGTGGAGCTGGACACCGT 797
Db 947 CCGCGCTACCTGTGGAGCTGAGACATGTGAAGCGGCTGGGCTCCACGACATCAT 1006

QY 798 C---ATCCGCCCAAGGGCAATTGAGGCAATGTGAACCGCATGGACATGAGGAGTGGAA 854
Db 1007 CGTTTCAAGCACTGGAAGATCAATGCGCATCTGTAACGGCATCGACACCAAGGAGTGGAA 1066
QY 855 CCCAAGACCGCAAGTTCCTGCTGCTGCGCCCTTACGACAGAACACAGG-----T 902
Db 1067 CCCAAGTGGACGTGCACTGCGTGGTGACGGCTACACCAACTCTCCCTCGAGACACT 1126
QY 903 CTACGCCGCAAGGCGCGCCGCAAGGAGGCGCTGAGGCGCGAGCTGGGCTGCCCTGTGTA 962
Db 1127 CGACGCTGGAAGCGGCACTGCAAGGCGCCCTGAGCGGAGCTGGGCTGGAAGTGG 1186
QY 963 CCCCACCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
Db 1187 CGACGACGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1246
QY 1023 CATCTGCGCGCCCTGCCCCAAGATCTGCGCACCCCAAGGTGCAAGCTCCATCTGCG 1082
Db 1247 CATCGGGGACGCGATGCGCTGGATC---GCGGGGAGGAGCTGAGCTGCTGCTGCTG 1303
QY 1083 TACCGCAAGGCGCGCTTACGAGAAGCTGTGAACGCACTCGGCACCAAGTACAAGGGCG 1142
Db 1304 CACCGGCGCGCGCTGGAACGATGCTGAGCACTTGGAGCGGAGCATCCCAACAA 1363
QY 1143 CCGCAAGGCGCTGCTGCAAGTCTTGGCGCCCTTGGCGCACATGCTCACCGCGCGCGCA 1202
Db 1364 GGTGCGCGGCTGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1423
QY 1203 CTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1262
Db 1424 CCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1483
QY 1263 CTACGCTACCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1322
Db 1484 ATACGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1543
QY 1323 CCGTACCGGCTTCCACATGGCGCCCTGAACCCCGACAGCTGGAGGAGCTGACCGCG 1382
Db 1544 CGACCGCTTACGCGACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1603
QY 1383 CGCCTTGGCGCGCACCGTGGCGCTGCCAGCGAGTGTGCTGGCGGCGCGCTACCCCG 1442
Db 1604 GATCGAGGCGCTCAGGCACTGCTGACAGTACCGGAACCTACGAGAGAGCTGGAGAG 1663
QY 1443 GATGCTGGCAACTGCATCAGCCAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1502
Db 1664 TCTCCAGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1723
QY 1503 GGGCTGCTGGAGGAGTGGTGTAC 1527
Db 1724 GGACGCTCTGTCAAGGCCAAGTAC 1748

RESULT 13

US-08-941-445A-8
; Sequence 8, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanning
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESS: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

Best Local Similarity 53.28: Pred. No. 9.4e-22:


```
QY 923 CCAAGAGCCCTGAGAGCCGAGCTGGGCTGCTGTGTGACCCACCAGCCGCCCTGTTCG 982
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1060 GTAAAGCTGAATTGAGAGGAGTTGGTTTACCTGTAGGGAGGATGTTCTCTCATTTG 1119
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 983 CCTTCATCGGCGCTGGAGGAGCAGAGGCTGTGACATCATCTGCGCGCTGCCCA 1042
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1120 GCTTTATTGGAAGACTGATTTACCAGAAAGCATGATCTCAATAAATGGCCATTCAG 1179
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1043 AGATCTGGCCACCCCGCAGGTGCGATCGCATCTGGGTACCGGCAAGCGCGCTAGG 1102
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1180 AGCTCATG---AGGGAGGACGTGCAATTGTCATGCTTGGATCTGGGGATCCAAATTTTG 1236
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1103 AGAAGCTGTGACGCCATCGCCACCAAGTACAAAGGCCCGCCCAAGGGCGTGGTCAAGT 1162
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1237 AAGGCTGGAAGATCTACCGAGTCGAGTTACAAGGATAAATTCGTTGGATGGTGGAT 1296
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1163 TCTCGGCGCCCTGGCGCATGCTCACCGCGCGCGGCGACTTCATGCTGGTGGCTCGC 1222
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1297 TTAGTGTTCAGTTTCCACAGAAATACTGCAGTTGCGATATATTGTAATGCCATCGA 1356
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1223 GCTTCGAGCCCTGGCGGCTGATCCAGCTGCACGCCATGCACTACGTTACCGTGGCGGTGG 1282
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1357 GATTTCGAACCTTGGGCTCTTAATCAGCTATATGCTATGCAATATGTACAGTTCCTGTTAG 1416
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1283 TAGCCTCGACCGCGGCGCTGGTCTGACACCGTCAAG 1317
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1417 TTCATGGAACTGGGGGCGCTCCGAGACACAGTCGAG 1451
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: June 3, 2003, 15:32:10
Job time : 69.3196 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 21:49:59 ; Search time 144.611 Seconds
(without alignments)
15416.986 Million cell updates/sec

Title: US-09-980-771A-8

Perfect score: 1593

Sequence: 1 gcgtgacatcgatggt.....ccgtgacgagagatcccc 1593

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 870385 seqs, 699768693 residues

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	474.8	29.8	2267	9	US-09-961-077-25
2	314.6	19.7	2399	9	US-10-138-075-1
3	275.2	17.3	2179	9	US-10-138-075-3
4	218.2	13.7	2825	10	US-09-952-677-5
5	118.2	7.4	2239	10	US-09-952-677-1
6	94.2	5.9	790	10	US-09-966-881-44
7	83.2	5.2	1427	10	US-09-974-300-757
8	82.2	5.2	805	10	US-09-966-881-43
9	76.4	4.8	813	10	US-09-966-881-45
10	76	4.8	1318	9	US-09-934-900-15
11	74.4	4.7	2712	10	US-09-748-033-4
12	72.2	4.5	299	10	US-09-294-093B-3520
13	72.2	4.5	824	10	US-09-894-633A-83
14	70.8	4.4	1415	9	US-09-934-900-11
15	70	4.4	3624	9	US-09-988-462-6
16	69.6	4.4	1896	9	US-10-124-880-15
17	69.6	4.4	3468	9	US-09-988-462-2
18	68.2	4.3	2681	9	US-10-232-563-1
19	68	4.3	804	10	US-09-966-881-40

20	66.8	4.2	4810	9	US-10-138-221-6
21	66.6	4.2	826	10	US-09-966-881-41
22	65.2	4.1	1929	9	US-09-899-642-1
23	65.2	4.1	2109	10	US-09-815-242-7761
24	65.2	4.1	88421	9	US-09-976-059-1
25	63.6	4.0	1059	9	US-10-232-563-4
26	62.6	3.9	828	10	US-09-815-242-7971
27	62.6	3.9	13842	9	US-09-860-846-30
28	62.6	3.9	13842	9	US-09-988-384B-30
29	62.6	3.9	13842	9	US-09-836-821-30
30	62.6	3.9	13842	10	US-09-861-289-30
31	62.6	3.9	36778	9	US-09-860-846-5
32	62.6	3.9	36778	10	US-09-836-821-5
33	62.6	3.9	36778	10	US-09-861-289-5
34	62.6	3.9	37948	9	US-09-988-384B-5
35	62	3.9	660	10	US-09-815-242-7875
36	62	3.9	1185	10	US-09-887-576-784
37	61.4	3.9	2442	10	US-09-815-242-4030
38	61.4	3.9	15872	9	US-09-860-846-1
39	61.4	3.9	15872	9	US-09-988-384B-1
40	61.4	3.9	15872	9	US-09-836-821-1
41	61.4	3.9	15872	10	US-09-861-289-1
42	61.4	3.9	15952	9	US-10-171-311-51
43	61	3.8	5224	9	US-09-373-658-32
44	60.8	3.8	1248	9	US-09-860-846-7
45	60.8	3.8	1248	9	US-09-988-384B-7

ALIGNMENTS

RESULT 1

US-09-961-077-25
; Sequence 25, Application US/09961077
; Publication No. US20030014775A1

GENERAL INFORMATION:

APPLICANT: Zwick, Michael G.
Edlington, Brent E.
McSwiggen, James A.
Merlo, Patricia Ann Owens
Guo, Lining
Skokut, Thomas A.
Young, Scott A.
Folkerts, Otto
Merlo, Donald J.

TITLE OF INVENTION: COMPOSITION AND METHODS FOR
MODULATION OF GENE EXPRESSION
IN PLANTS

NUMBER OF SEQUENCES: 1263

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
Storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/961,077

FILING DATE: 21-Sep-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/679,645

FILING DATE: July 12, 1996

APPLICATION NUMBER: 60/001,135

FILING DATE: July 13, 1995

APPLICATION NUMBER: 08/300,726

Sequence 6, Appli
Sequence 41, Appli
Sequence 1, Appli
Sequence 7761, Ap
Sequence 1, Appli
Sequence 4, Appli
Sequence 7971, Ap
Sequence 30, Appli
Sequence 30, Appli
Sequence 30, Appli
Sequence 30, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 7875, Ap
Sequence 784, App
Sequence 4030, Ap
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 51, Appli
Sequence 32, Appli
Sequence 7, Appli
Sequence 7, Appli

FILING DATE: September 2, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 219/247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 2267 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 25:

US-09-961-077-25

Query Match 29.8%; Score 474.8; DB 9; Length 2267;
Best Local Similarity 59.9%; Pred. No. 2.7e-105;
Matches 978; Conservative 0; Mismatches 557; Indels 99; Gaps 7;

5 TGGACATCGTATGGTGTCTGCTGAGGTGCGCCCTTGGTCCAAAGACGGCGGCTTGGCG 64
1147 CCGAGGAGCTCATCTCCGGCATCGCAGGGCTGCGAGCTGACAAACATCATGCGCTCA 1206

809 AGGCATTAGGGGATTTGTGAACGGCATGGACATTTAGAGAGTGGACCCCAAGACCCGACA 868
1207 CCGGCATCACGGCATCTGTAACGGCATGGACATTTAGAGAGTGGACCCCAAGACCCGACA 1266

869 AGTTCTCTGTCGCGCTTACGACCAAGACAGCTTACGCGCGGCAAGCGCGCGCAAGG 928
1267 AGTACATCGCTGAAGTACAGCTGTGCGCGCTGAGGCGCAAGCGCTGACACAGG 1326

929 AGGCCCTGACGCGGAGCTGGGCGCTTGTGTGAGCCCAACCGCCCGCTTGTGCGCTTCA 988
1327 AGCGCTGACGCGGAGGTGCGGCTTCCGGTGGACCGAATCATCCGCTGGTGGGCTTCA 1386

989 TCGCGCGCTGAGGAGCAGAGGCTGTGGACATCATCTGCGCGCTGCGCCCAAGATCC 1048
1387 TCGCGAGCTGAGAGCAGAGGAGCCCGAGCTCATGGGCGCCCAATCCCGCAGCTCA 1446

1049 TGGCACCCCAAG---GTGAGATCGCCATCTCTGGTACCGGCAAGCGCGCTTACGAGA 1105
1447 TGGACATGGTGGAGACGTGCAGATCTCTCTGGCACCGGCAAGAAAGATTCGAGC 1506

1106 AGCTGGTGAACGCCATCGGACCAAGTACAGGGCGCGCCCAAGGGCGTGGTCAAGTTCT 1165
1507 GCATCTCATGAGCCCGAGGAGAAATTCCTCAGGCAAGGTGCGCGCTGCTCAAGTTCA 1566

1166 CGGCGCGCTGCGCACATGCTACCGCGCGGCGGCTTCTATGCTGCTGCTCGCTCGCT 1225
1567 ACGGCGCGCTGCGCACCATCATATGCGCGCGGCGGCGGCTGCTGCGCTGCTGCGCT 1626

1226 TCGAGCCCTGCGCGCTGATCCAGCTGCGCGCTATGCTGCTGCTGCTGCTGCTGCTG 1285
1627 TCGAGCCCTGCGCGCTTATCCAGCTGCGAGGCGATGCTATGCTGCTGCTGCTGCTG 1686

1286 COTCCACCGCGCGCTGCTGCGCACCGCTCAAGAGGCGCTGCTGCTGCTGCTGCTGCTG 1345
1687 COTCCACCGCTGCTGCTGCGCACCGCTCAAGAGGCAAGGCTGCTGCTGCTGCTGCTG 1746

1346 COTCTCA-----ACCGCGACAGCTGAGGAGCTGAGCGCGCTGCTGCTGCTGCTGCTG 1399
1747 GCCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1806

1400 TCGCGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1459
1807 TCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1866

1460 TCGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1519
1867 TCGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1926

1520 TGGTGTACGCGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1579
1927 TCGGGTGC---GCGCGAGCGCGGCTGAGGCGGAGGAGTGGAGGAGTGGAGTGGAGT 1983

1580 CCGAGAGATCC 1593
1984 AGGAGAACGTGGCC 1997

RESULT 2

US-10-138-075-1
; Sequence 1, Application US/10138075
; Publication No. US20030087369A1
; GENERAL INFORMATION:
; APPLICANT: Broglie, Karen E.
; APPLICANT: Butler, Karlene H.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Orozco, Emil M.
; TITLE OF INVENTION: Granule-Bound Starch Synthase

; LENGTH: 2179

; TYPE: DNA
; ORGANISM: Glycine max
US-10-138-075-3

Query Match 17.3%; Score 275.2; DB 9; Length 2179;
Best local similarity 52.7%; Pred. No. 3e-57;
Matches 820; Conservative 0; Mismatches 628; Indels 108; Gaps 6;

```
QY 11 TCGTGATGGTGTGCTGAGTCCGCCCTTGGTCCAGACGGCGCGCTGGCGCATGGA 70
DB 323 TCATAATATAGGAACCGAGGTGGCTCCATGGTCAAAACCTGGGGTGGGATGTC 382
QY 71 CTGTGGCGCTGCTATTGAGTGGTCAAGCGCGCCACCGCTCATGACCAATTCGCCCTC 130
DB 383 TTGGAGGCTPACCAACCGCATTGGCAGTTTGGGCATCGAGTAATGACTATTGTGCGC 442
QY 131 GCTACGACGATAGCTGACGCTCGGGACACCTCGGTGGTGGTGGACATCATG----- 184
DB 443 GCTATGACCAATCAAAAGATGATGGATGATCAAGTGTGTAATGAGGTGAATAGGAG 502
QY 185 -----GCGAGAAGTCCGCTACTTCCACTCCATCAAGAAGGGGTGCAACCGGTGGGA 238
DB 503 ATAGAACAGAAAGTTTGGCTTCTTCCATGTTATAGAGGGGAGTTCATGCTGTCTTG 562
QY 239 TTGACCAACCCCTGCTTCTGCGCAAGGTCTGGGCAAGACCGGCTCCAAGCTGTACGGCC 298
DB 563 TGGATCACCCCTTGGTTCTTGAAGAGTATGGGGAAACCTGGACAAAACCTTTATGGAC 622
QY 299 CCGCTCGGGCGCTGACTACCTGGACACACACACAGCGGTTCGCCCTGTTCGAAGGCG 358
DB 623 CAACTACTGGAATGATTACGAAGACACCAACTGCGTTTAGCCTTTTGGCAGGCTG 682
QY 359 CTATTGAGGCTGCGCGGTGCTGCCCTTCGCGCCC-----G 394
DB 683 CTTTGGAGCCCAAGGGTTCAGTCTTATTCAGTAATATTTCTCTGACCAATG 742
QY 395 GCGAGGACTCGCTTCTGTCGCAAGCTGCGACTCGCCCTCGCTGGTGGCGCTGCTGTA 454
DB 743 GTGAAGATGTCATTTTGTCCCAATGATTGGCACACTGCGCTTATCCCTGCTACTTGA 802
QY 455 AGGAGGATACAGCCCAAGGCGAGTTCACCAAGGCAAGCTGCTGCTGCTATCCACA 514
DB 803 AAGTATGTACAGTCAAGGGGCACTATAGCAATGCGCGGTGTTTTTGTATCCACA 862
QY 515 ACATCGCTTCCAGGCGCGATGTGGGAGGAGGTTTCAAGGACACGAGTGCCTCCAGG 574
DB 863 ACATGCTTACCAAGGAAGTTGATTCGCGGACTTCTCACTTAACTCTCCAGACC 922
QY 575 CGGCTTTGACAAGTGGCTTCTCGGAGCGGTATGCCAAGTTTACACTGAGGCCACCC 634
DB 923 AATTAAGAGCTCTTTCACCTTATTTGATGGCATG----- 958
QY 635 CCATGGAGGAGGACGAGAAGCCCGCTGACGGGAAGACCTACAGAAGATCAACTGGC 694
DB 959 -----TTAACCACTGGTGGGAAGGAATCAATTTGTT 991
QY 695 TGAAGGTGGCATTATCCCGCGCAAGCTGGTGAAGTGTGCGCCCACTACGGACCG 754
DB 992 TGAAGCTGGACTTATAGAAATCATGTTTGTGATAACCGTTAGTCCAAACTATGCTAAAG 1051
QY 755 AGATCGCTGCGATGCCCGCGGTGGAGCTGGACACCGCTCATCGCGCCAAG--- 811
DB 1052 AACTGGTGTGAGTCCAGACAAAGAGGTGGAATTTGGCAACATCATTCGCAAAATTTGATG 1111
QY 812 -----GCATTGAGGCAATTGGAACGGCATGGACATTCAGGAGTGGAAACCCCAAGA 862
DB 1112 ATGATGCTGTTTGGTGGATTTGATGCAATGGATGTTTCAGAGTGGATCAACCA 1171
QY 863 CGCAAGTTCCTGTCTGCGCCCTACGACCAAGACAGCTGTACCGCGCAAGCGCGCG 922
DB 1172 CTGCAAAATATATAGCTGCAAAATATGATTTTCAACAGTATTGGAAGCAAGGCTCTTC 1231
QY 923 CCAAGGAGGCGCTGACGCGAGCTGGCGCTGCTGTGACCCCAACCGCCCTCTGTCG 982
```

```
DB 1232 TGAAGAAGCCCTCCAGCAGAAGTTGGATTGCCAGTCCACAGAAATATTCCTCTCATG 1291
QY 983 CTTTCATCGCGCGCTGGAGGAGAGAGGTGTGGACATCATCTGCGCGCCCTGCCCA 1042
DB 1292 GTTTCATTGGTAGGCTTGAAGAGCAAAAGGTTCTGATATCTTTCGCAAGGCAATCCCC 1351
QY 1043 AGATCTCGCCACCCCAAGGTGCAGATCGGCATCCTGGGTACCGCGCAAGCGCGCTACG 1102
DB 1352 AATTAT---CAAGCAGATGTTCAAGTTGGTAGCCCTAGGAACAGAAAAACAATGG 1408
QY 1103 AGAAGCTGGTGAACGCCATCGGCACCAAGTACAGAGGCGCGCCCAAGGCGGTGGTCAAGT 1162
DB 1409 AAAAGCAGCTTGAGGAAGTGAATAATATCATACCTGTATAGGCCAGAGAGGTGCAAAAT 1468
QY 1163 TCTCGCGCCCTGCGCACATGCTCACCGCGCGCGGCTTCATGCTGTGCGCCCTCGC 1222
DB 1469 TCAATGTTCCCTAGCCACATGATATAGCTGGAGCTGATTTATATGTTGTTCTAGCA 1528
QY 1223 GCTTCGAGCCCTGCGGCTGATCCAGCTGCAAGCCATGACCTACCTACCGTACCGCTGG 1282
DB 1529 GATTGAGCCCTGTGCTCATTCAGTTACAGCTATCGCTATGATCTGTACCAATG 1588
QY 1283 TAGCCTCACCGCGCGCTGCTGACACCGTCAAGAGGCGGTACCGGCTTCCACATGG 1342
DB 1589 TTGCTCAACAGGTGATGTTGACACTGTCAAGAGGCTTCAAGAGGCTTCAAGATTCAGATGG 1648
QY 1343 GCGCCCTGA-----ACCCGCAAGCTGGAGGCTGACCGCGCGGCTTCCGCGCGCA 1396
DB 1649 GTGCTTCAATGTTGAATGTGATGCTGTGGATCCGCTGATGTGATGCTATATCAAGA 1708
QY 1397 CCGTGGCGCGTCCAGCGAGGTGTTTGGCGGGCGCGCTACCGCGAGATGGTGGCGCACT 1456
DB 1709 CTGTCAAAAGGCGCTTGCAGTCTATGGAACITCCAGCTTTTACAGAAATTTATCAAGACT 1768
QY 1457 GCATCAGCAGGACCTGCTTCGTCCAGCGCGCCAGAGGCGGCGCTGCTG 1512
DB 1769 GCATGCGCTCAAGATCTTTCATGGAAGGCGCTGCTAAGAGTGGGGAAGTGCCTG 1824
```

RESULT 4

US-09-952-677-5

; Sequence 5, Application US/09952677

; Patent No. US20020138876A1

; GENERAL INFORMATION:

; APPLICANT: Block, Martina

; Lorz, Horst

; Lutticke, Stephanie

; Walter, Lennart

; Frohberg, Claus

; Kossmann, Jens

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES

; FROM WHEAT WHICH ARE INVOLVED IN STARCH

; SYNTHESIS

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave

; STREET: 1251 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: United States of America

; ZIP: 10020

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA: US/09/952,677

; FILING DATE: 14-Sep-2001

; PRIOR APPLICATION NUMBER: 09/196,390

; FILING DATE: 19-No. US20020138876A1-1998

		Query Match	13.7%	Score	218.2;	DB	10;	Length	2825;	
		Best local similarity	50.1%;	Pred.	No. 1.6e-43;					
		Matches	769;	Conservative	0;	Mismatches	688;	Indels	78; Gaps	6;
QY	5	TGACATCGTGTGGCTTGCTGCTAGTCGCCCCCTTGGTCCAAAGCGGGCGCCTGGCGG	64							
Db	1084	TGAACATGGTGCFCGTGGCTGCTGAGTTCTCCCTGGTCAAACAGSGTGTCTGGAG	1143							
QY	65	ATGTACTGGTGGCCCTATTGACCTGGTCAAGCGCGCCACCGCGTCATCACCATTTG	124							
Db	1144	ATGTTGGGCTCTGCCCCAAGCWTGGCAAGAAGAGACATCGTGTATTTGTTG	1203							
QY	125	CCCCTCGCTACGACCAGTAGCTGAGCGCTGGGACACCTCGGTGGTCTGGACATCATGG	184							
Db	1204	TACCAAGGTATGGGACTATGAAGAAGCCTACGATCGGAGTCCGAAAATACTACAAG	1263							
QY	185	GCGAGAAGTCCGCTACTTCCACTCCATCAAGAAGGCGGTGCACCGCGTGTGGATTGACC	244							
Db	1264	CTG-----CTGGACAGGATATGGAAGTAATTAITTCATCGTTATATCGA	1309							
QY	245	ACCCCTGGTCTCTGGCCAAGTCTGGGGCAAGACCGCTCCAACTGTAGGCGCCCGCGT	304							
Db	1310	TGGAGTTGATTTTGTGTTCAATGACGCTCCTCTCTTCCGAC-ACCGTCAGGAAGACATTT	1368							
QY	305	CGGGCGCTCACTACCTGGACACACCAAGCGTTCGCCCTGTCTCAAGCGCGCTATTTC	364							
Db	1369	ATUGGGGCACACACAGGAAATATGAACGCATGATTTGTTCTCAAGGCGCTGTTCG	1428							
QY	365	AGGCTCCGCGGTGCTGTCCTCTTGGCCCGGCGGACGACTGCG-----TCTTCG	412							
Db	1429	AGGTTCCATGGCAGGTTCCATGGGCGGTGTCCTTATGGGATGGAATCTGGTGTTTA	1488							
QY	413	TGCCAACGACTGGCACTCCGCCCTGGTCCCGTCTGCTGAAGCAGGASTACCA GCCCA	472							
Db	1489	TTGCAATGATTTGGCACACGGCACTCCTGCCCTGTCTATCTGAAGCATATTACAGGGACC	1548							
QY	473	AGGGCCAGTTACCAAGGCCAAGTCGGTCTGTGCTATNCCACAACATPCGCTTCCAGGGCC	532							

Sequence 1, Application US/0952677
Patent No. US20020138876A1

GENERAL INFORMATION:

APPLICANT: Block, Martina

Lorz, Horst

Luticke, Stephanie

Walter, Lennart

Prober, Claus

Kossmann, Jens

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
FROM WHEAT WHICH ARE INVOLVED IN STARCH
SYNTHESIS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave

STREET: 1251 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10020

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/952,677

FILING DATE: 14-Sep-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/196,390

FILING DATE: 19-NO. US20020138876A1-1998

APPLICATION NUMBER: DE 196 21 588.9

FILING DATE: 29-MAY-1996

APPLICATION NUMBER: DE 196 36 917.7

FILING DATE: 11-SEP-1996

APPLICATION NUMBER: PCT/EP97/02793

FILING DATE: 28-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Haley, Jr., James F.

REGISTRATION NUMBER: 27,794

REFERENCE/DOCKET NUMBER: AGREVO-9

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 596-9000

TELEFAX: (212) 596-9090

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2239 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Triticum aestivum L.

STRAIN: cv. Florida

HAPLOTYPE: ca. 21 d Caryopses

IMMEDIATE SOURCE:

LIBRARY: cDNA library in pBluescript sk (-)

CLONE: TaSSS

FEATURE:

NAME/KEY: CDS

LOCATION: 3..2017

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-952-677-1

Query Match 7.4%; Score 118.2; DB 10; Length 2239;
Best Local Similarity 52.4%; Pred. No. 1.9e-19;
Matches 333; Conservative 0; Mismatches 293; Indels 9; Gaps 3;
US-09-952-677-1

QY 686 TCAACTGGCTGAAGGTGGCAATATCGCGCGGACAAAGCTGTGACTGTGTGCGCCCAACT 745
Db 823 TTAACVTTTGAAGAGGACAGTGTGACAGACAGATCGGATTTGTACCCCTCAGTCAGGGTT 882

QY 746 ACGGACCGAGATC---GCTCCGATGCCCGCGGCTGTGGAGCTGGACACCGTCATCC 802
Db 883 ATTCATGGGAGGTCACTCACTGCTGAAGGTGGACAGGCGCTCAATGAGCTCTTAAGCTCCC 942
QY 803 GCGCAAGGCAATTGAGGCAATTGTGAACGCGATGGACATTTGAGGAGTGGACCCCAAGA 862
Db 943 GAAAAGTGATTTGAATGGAATTTAAATGAATTTGACATTAATGATTTGAAATGGAATTTG 1002
QY 863 CCGACAAGTCTCTGCTCGCGCTACGACAGACAGAGCTGTACGCCGCGAAGCGCGCG 922
Db 1003 CAGCAAGTGTCTCCCTCATCTATTCTGTGATGACTCT---CTGGAAGGCAAT 1059
QY 923 CCAAGGAGCGCTCGAGCGCGAGCTGGCGCTGCTGTGGACCCCAAGCGCGCGCTGTTCG 982
Db 1060 GTAAAGCTGAATTCAGAGAGGAGTTGGGTTTACCTGTAAAGGAGGATTTCTCTGATTG 1119
QY 983 CTTTCATCGCGCGCTGGAGGAGAGGAGTGTGACATCATCTGCGCGCGCTGCGCA 1042
Db 1120 GCTTTATTGGAAGACTGGATTACCAGAAAGGCAATGCTCTATTAAATGGCCATTCAG 1179
QY 1043 AGATCTGGCCACCCCAAGTGCAGATCGCATCTCTGGGTACCGGCAAGCGCGCTACG 1102
Db 1180 AGCICATG---AGGAGGAGCGTGCAATTGTGATCTGGGATCCCAATTTTG 1236
QY 1103 AGAAGCTGTGAACGCGCTATCGGCACCAAGTACAGGCGCGCGCGCGCGCTGGTCAAGT 1162
Db 1237 AAGCTGTGATGAGATCTACCGAGTGCAGTTACAGGATAAATTCCTGATGGTGGAT 1296
QY 1163 TCTCGCGCGCTGGCGCATGTCACCGCGCGCGCTTCATCTGCTGGTCCCTCCG 1222
Db 1297 TTAGTGTTCCTAGCTTTCCACAGAAATACATGCAAGTGTGATATATTGTAATGCCATCA 1356
QY 1223 GCTTCGAGCGCTCGCGCTGATCCAGCTGCAGCGCATGCACTACCGTACCGTCCCGTGG 1282
Db 1357 GATTTGAACCTTGGCGTCTTAATAGCTATATGCTATGCAATATGCTACAGTCTCTGTAG 1416
QY 1283 TAGCCTCCACCGCGCGCTGGTGCACACCGTCAAG 1317
Db 1417 TTCTATGAAGTGGGCGCTCCGAGACACAGTCGAG 1451

RESULT 6

US-09-966-881-44
Sequence 44, Application US/09966881
Patent No. US20020120960A1

GENERAL INFORMATION:

APPLICANT: Seymour, Graham

Bird, Colin

Medina-Suarez, Rosybel

TITLE OF INVENTION: Genetic control of Fruit Ripening

NUMBER OF SEQUENCES: 57

CORRESPONDENCE ADDRESS:

ADDRESSER: Zeneca Ag Products Inc.

STREET: 1800 Concord Pike

CITY: Wilmington

STATE: DE

COUNTRY: USA

ZIP: 19850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/966,881

FILING DATE: 28-Sep-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/242,850

FILING DATE: 29-Mar-1999

APPLICATION NUMBER: GB 9618862.3

FILING DATE: 10-SEP-1996

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/966,881
FILING DATE: 28-Sep-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/242,860
FILING DATE: 29-Mar-1999
APPLICATION NUMBER: GB 9618862.8
FILING DATE: 10-SEP-1996
APPLICATION NUMBER: GB 9708366.1
FILING DATE: 25-APR-1997
APPLICATION NUMBER: PCT/GB97/02424
FILING DATE: 08-SEP-1997

ATTORNEY/AGENT INFORMATION:
NAME: Hohenschutz, Liza D.
REGISTRATION NUMBER: 33,712
REFERENCE/DOCKET NUMBER: SEE 50183/UST
TELEPHONE: (302) 886-1699

SEQUENCE CHARACTERISTICS:
LENGTH: 805 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: cdna
IMMEDIATE SOURCE:
CLONE: U-D66
SEQUENCE DESCRIPTION: SEQ ID NO: 43:

US-09-966-881-43

Query Match 5.2%; Score 82.2; DB 10; Length 805;

Best Local Similarity 57.8%; Pred. No. 8.8e-11;

Matches 170; Conservative 0; Mismatches 111; Indels 13; Gaps 1;

QY 1 CGCTGGACATCGTGTGCTGAGTGGCCCTTGCTCCAGACGGGGCGCTG 60
DB 310 GGGATGACATAGTCTTTGTCGGCGCTGAGATGGCTCCTGGATTAACCGGAGACTC 369
QY 61 GCGGATGACTGTGGCGCTGCTATGAGCTGTGTCAGCGCGCCACCGGCTCATGACC 120
DB 370 GGTGATGTCCTGGAGGACTGCCACCGCCATGGCTGCAATGGACACACAGTGATGACT 429
QY 121 ATTGCCCTCGCTAGCACCAGTACGCTGACGCTGGGACACCTCGGTGTCG----- 172
DB 430 ATAGCTCCACGCTATGATCAGTACAAGGATGGTGGGATACAAATGTCCTGGCTGANTTA 489
QY 173 -----TGGACATCATGGCGGAGAAGTCCGCTACTTCCACTCCATCAAGAAGGGCGTGCA 227
DB 490 AAANTTGAATGAATTTGAAGAGANTCCGCTCTTCCACTGCTATAAANAGATTGA 549
QY 228 CCGCGTGTGATGATACACCCCTGTTCTGCCCAAGTCTGGGGCAAGACCGG 281
DB 550 CAGGTTTTCATGATCATCTTTGTTCTTGAAAAAGTGTGGGAAAAAAGTGG 603

RESULT 9

US-09-966-881-45/c

; Sequence 45, Application US/09966881

; Patent No. US20020120960A1

; GENERAL INFORMATION:

; APPLICANT: Seymour, Graham

; Bird, Colin

; Medina-Suarez, Rosybel

; TITLE OF INVENTION: Genetic control of Fruit Ripening

; NUMBER OF SEQUENCES: 57

; CORRESPONDENCE ADDRESS:

; ADDRESSER: Zeneca Ag products Inc.

; STREET: 1800 Concord Pike

CITY: Wilmington
STATE: DE
COUNTRY: USA
ZIP: 19850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/966,881
FILING DATE: 28-Sep-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/242,860
FILING DATE: 29-Mar-1999
APPLICATION NUMBER: GB 9618862.8
FILING DATE: 10-SEP-1996
APPLICATION NUMBER: GB 9708366.1
FILING DATE: 25-APR-1997
APPLICATION NUMBER: PCT/GB97/02424
FILING DATE: 08-SEP-1997

ATTORNEY/AGENT INFORMATION:

NAME: Hohenschutz, Liza D.
REGISTRATION NUMBER: 33,712
REFERENCE/DOCKET NUMBER: SEE 50183/UST
TELEPHONE: (302) 886-1699

INFORMATION FOR SEQ ID NO: 45:

SEQUENCE CHARACTERISTICS:
LENGTH: 813 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: cdna
IMMEDIATE SOURCE:
CLONE: U-D112
SEQUENCE DESCRIPTION: SEQ ID NO: 45:

US-09-966-881-45

Query Match 4.8%; Score 76.4; DB 10; Length 813;

Best Local Similarity 54.8%; Pred. No. 2.2e-09;

Matches 194; Conservative 0; Mismatches 153; Indels 7; Gaps 2;

QY 1157 TCAAGTTCTCGCGCCCTCGGCACATGCTCACCGCGCGCGGCGGCACTTCATGCTGGTGC 1216
DB 535 TGAAGTTTAATGCACCTTTGGTGCATGAAATCATGCTGGAGCAAGATCTTCTGNTCTTA 476
QY 1217 CCTCGGCTTCAGCCCTCGCGCCCTGATCCAGC-TGCAGCCCATGCACCTACCGTACCGTG 1275
DB 475 CCAGCAGATTCCAACTTTGGCCCTATCCAGCTTTCAGGGCATCGGATATGGAATCCC 416
QY 1276 CCGCTGGTAGCTCCACCGCGGCTGTGTCGACACCGTCAAGGAGGGCGTCAACCGGCTTC 1335
DB 415 CCCATGTGCGCAACAACTGCTGACTGCTGACACTGCTCATAGAAGGCATCACAGGTTT 356
QY 1336 CACATGGGCGCCCTGAACCCCGACAAAGCTGGAGGCTGACCGCGACGCCCTGGCCGCC 1395
DB 355 CATATGGGTCCCTTCAATCCTGATTGTGATGTTGTGATTAAGATGATGATGATGATGATG 296
QY 1396 -----ACCGTGGCGGCTCCAGCGAGGTGTTTGGGGGGCGGCTACCCCGCATGGTG 1449
DB 295 ATTACACAGATGAAAAGGGCCCTTAAAGTTTACGGCACACCTGCACTTCTGCTGAGATGATA 236
QY 1450 GCGAACTGCATCAGCAGGACCTGCTGCTGCCAAGCGCGCGCGCGCGCGAGAGTGGAG 1503
DB 235 CAGAACTGCATGACTCAAGACCTCTCTCTGGAAGGACCTGCTAAGAAGTGGGAG 182

RESULT 10

US-09-934-900-15

; Sequence 15, Application US/09934900

; Publication No. US20030054521A1

Db 1455 CGGTCCAGGCGCGACACACCGCGACAGACCGCGCGGTACGCGCCCGCGTCAAGGTCA 1514
QY 1166 CGGCGCCCTGGCGCATCTACCGCGCGCGCGGCGGACTCATCTGTGTGCGCTCGGCT 1225
Db 1515 CCACCCCGCGGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1568
QY 1226 TCAGCGCTGCGGCGCTGATCCAGCTGACGCTGACGCGCATCTACGCTACGCTGCGCGGTAG 1285
Db 1569 TCGCTACTTACCGAGTGGGCGGTCTACGCGCGCACTACCACTGCAAAACCTGGTGA 1628
QY 1286 CCTCCAGCGGCGCTGTCACACGCTCAGGAGGCGTCAACGCGTTCACATGCGCG 1345
Db 1629 CCTCGGCTCGCGGAGAGTACCCACCACTACTCTCTCGCAACCTCCAGGCG 1698
QY 1346 CCCTGAACCGCGCAAGCTGACGAGGTGACGCGGCGGCGGCGGCGGCGGCGGCGGCG 1405
Db 1589 GC---AGTGACCATCGGTGACAGCTTCGCGCGCTACGACAGCGGTACCGCGCGG 1745
QY 1406 GTGCCAGCGAGGTGTTTCGCGGCGCGCGCTACCCCGGAGAGTGGTSCCACTGCATAGCC 1465
Db 1746 AGTCGTCGAGCGGCTGCGGACACCTTGGGACGAGCGCTGCGCGGCACTTCAACACG 1805
QY 1466 AGGACCTGCTCTGTCTCAAGCGCGCGGCGGAGAGTGGGAGGCGCTGCTGAGAGTGGTGT 1525
Db 1806 TCCGAAGCTCAAGGCGCAAGTACCCGACATCAAGGTCCTCTGCTTCCTTCGCGGCGTGA 1865
QY 1526 ACGGCAAGGCGGCGGTGCGCCACCGCGCAAGAGGA 1559
Db 1866 CTGTGTCGCGGCTTCACCGACGCGGTGAAGA 1899

RESULT 12

US-09-294-093B-3520
; Sequence 3520, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 3520
; LENGTH: 299
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700380278H1
; NAME/KEY: unsure
; LOCATION: 118
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-3520

Query Match 4.5%; Score 72.2; DB 10; Length 299;
Best Local Similarity 60.1%; Pred. No. 2.2e-08;
Matches 119; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
QY 1130 AGTACAAGCGCGCGCGCGCGCGGTGTCAGTTCGCGCGCGCGCGCGCGCGCGCGCGCG 1189
Db 75 AATACCATGGCGGAGTGAAGTATGCTTAAGCTATGACGAGCGCCTGTCACATTTGATAT 134
QY 1190 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1249
Db 135 ATGCTGGCGGACAGCTCAATTCGTTCTTCACTTCGAACCTTGTGTCTAACACAGC 194
QY 1250 TGCAGCGCATGACATACGCTACCGTGCCTGGTGAAGCTCCACCGCGCGCGCGCGCGCG 1309

Db 195 TTATTCCTATGCGCTATGATCCATCCAGTTGTTTCGGAATACTGGAGGCGCTGTAGGACA 254
QY 1310 CCGTCAAGGAGGCGGTCA 1327
Db 255 CCGTTTTGATGTCGACA 272
RESULT 13
US-09-894-633A-83
; Sequence 83, Application US/09894633A
; Patent No. US20020124285A1
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy
; APPLICANT: Dubois, Patrice
; APPLICANT: Malven, Marianne
; APPLICANT: Masucci, James
; TITLE OF INVENTION: PLANT REGULATORY SEQUENCES FOR SELECTIVE CONTROL OF GENE EXPRESSION
; FILE REFERENCE: 38-21(15856)B
; CURRENT APPLICATION NUMBER: US/09/894,633A
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/214,357
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 09/894,633
; PRIOR FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 83
; LENGTH: 824
; TYPE: DNA
; ORGANISM: Zea mays
US-09-894-633A-83

Query Match 4.5%; Score 72.2; DB 10; Length 824;
Best Local Similarity 47.6%; Pred. No. 2.2e-08;
Matches 246; Conservative 0; Mismatches 268; Indels 3; Gaps 1;
QY 795 CGTATCCGCGCGCAAGGCGCATTTGAGGCGCATTTGACGCGCATTTGAGAGTGGAA 854
Db 132 CTTGCTCCGCGCTTTCTTCCAGAGTGTCTTCTCAGCGGGTCCGACGCGTGTGTGTAT 191
QY 855 CCCCAGACCGCAAGATTCTCTGCGCCCTACACACAGAGAGCGTGTACGCGGCA 914
Db 192 CGCTCCACCGAGTTCAGAGTCCGAGCACCACGCGGAGATCAACCACTCCCTCCCGG 251
QY 915 GCGCGCGCGCAAGAGGCGCTCGAGCGAGTGGCGGCGCTGTGGGCGCGCTGTGGACCCACCGCGG 974
Db 252 GGACGCTTCGACGCGCGTGTGCGGCGCAAGCTGCGCCCTGGAGTGGCGCGCGGCT 311
QY 975 CCGTTCGCTTCATCGCGCGCTGGAGGAGAGAGGCGTGTGGACATCATCTGCGCGC 1034
Db 312 GGTGT---CCTGCGCGCATCTCGACTGCGGCGTGTGGCGGCTGTATACCATACCG 368
QY 1035 CCGTCCCAAGATCTCTGGCGACCCCAAGGTGACATCGCATCTCTGGTGTACCGCAAGC 1094
Db 369 CGGCGCGCGGTACCGCGTTCCGCTGGGCGCGAGGAGTCTGCTGTGCTGCGCGCGCG 428
QY 1095 CGCCTACGAGAGTGGTGAACCGCATCGGCAACAGAGGAGGCGCGCGCGCGCGCG 1154
Db 429 GCGCGCGGTGGAGTGGCGCGCAACTTCACTGCGGCGCGCGCGCGCGCGCGCGCG 488
QY 1155 GGTCAAGTTCCTCGCGCGCGCTGGCGCATGTCTACCGCGCGCGCGCGCGCGCGCG 1214
Db 489 CGCAGAGGTTTACGGTTCAGAGTGTGTGGCGTGTGCGCGCGCGCGCGCGCGCGCG 548
QY 1215 GCGCTCGCGCTTCGAGCGCGCTCGGCGTGTATTCAGCTGTGACGCGCGCGCGCGCG 1274
Db 549 CTCCTCACTCAAGAGTGTGCGCGCGCGCGCGCTCACTCACTTCCGCAACAGCGCGG 608
QY 1275 GCGCGTGTAGCTTCCACCGCGCGCGCTGTGTGACAC 1311
Db 609 GGAGCAGTTCGAGCGCGCGCGCATGAACCGCGTCTACGCC 645

RESULT 14

US-09-934-900-11
; Sequence 11, Application US/09934900
; Publication No. US20030054521A1
; GENERAL INFORMATION:
; APPLICANT: Booth, Russ
; APPLICANT: Cahoon, Rebecca E
; APPLICANT: Hitz, William D
; APPLICANT: Kinney, Anthony
; APPLICANT: Yadav, Naren
; TITLE OF INVENTION: Nucleotide Sequences of a New Class of Diverged Delta-9 Stearoyl-
; FILE OF INVENTION: ACP Desaturase
; FILE REFERENCE: B01476 US NA
; CURRENT APPLICATION NUMBER: US/09/934,900
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 2001-08-22
; PRIOR FILING DATE: 2000-08-22
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 1415
; TYPE: DNA
; ORGANISM: Zea mays
US-09-934-900-11

Query Match 4.4%; Score 70.8; DB 9; Length 1415;

Best Local Similarity 44.8%; Pred. No. 4.9e-08;

Matches 316; Conservative 0; Mismatches 387; Indels 3; Gaps 1;

QY	712	GCCGCCGACAACTGGTGACTGTGTGCGCCAACTACGCGAGCGGAGATCGCTGCCGATGCC	771
DB	346	GCGGACTTCCTCCCGGACTCTGTCGAGATGTTCCGGCAGAGTCCGCGAGTCGCG	405
QY	772	GCCGGGGGTGTGGAGTGGACACCGTCTATCCGCCCAAGGCAITGAGGCAITGTGAAC	831
DB	406	GCCCGCGCGGGGGCTCCCGGACGAGTACTGCTGCTGCTGCTGGGGGACATGTCACG	465
QY	832	GGCATGACATTTAGGAGTGAACCCCAAGACCAAGTTCCTGTCTGCGCCCTTACGAC	891
DB	466	GAAGAGGGCTGCGCCAGTACCAAGACCATGATCAACAGCTCGACGGCTCCGCGACGAG	525
QY	892	CAGAACAGCTGTACGCGCGCAAGCGCGCCCAAGGAGGCGCTCGAGGCGGAGCTGGC	951
DB	526	ACCGGCGCCAGCACTCCCTGGGGTCTGGAGCGCGCTGGACCGCGGAGAGAAC	585
QY	952	CTGCCTGTGACCCCAACCGCCCCCTGTTCGCCCTTCATCGCCCGCTGGAGGAGCAGAAG	1011
DB	586	CGCCACGCGGACATCTCTCGGCAAGTACATGTACCTATCCGSCGCGTGCATGGCATG	645
QY	1012	GGTGTGACATATCTCTGGCGCGCTGCCCAAGATCTTGGCCACCCCGCAAGTGCAGATC	1071
DB	646	GTGAGAGACCGTCCAGTACCTCATCGGCTCCGCGATGATCCCGAACGAGAGAAC	705
QY	1072	GCCATCTGTGGTACCGGCAAGCGCGCTACGAGAAGCTGTGTAAGCCATCGCGACCAAG	1131
DB	706	CCGTACTTGGCTTCGTGTACACAGACTTCAGAGAGCGCGGACGGCGCTCTCCACGGC	765
QY	1132	TACAGGGCGCGCAAGGGCGTGGTCAAGTTCTTCGGCGCCCTTGGCGGCATCTCAACC	1191
DB	766	AACACCGCGGGCTCGCGCAGGGCGACGGGACGACGTCTTGGCGCGGCTTGGCGCAC	825
QY	1192	GCGGGCGCGACATTCATGCTGGTGGCTTCGGCTTCGAGCCCTCGGCGCTGATCCAGCTG	1251
DB	826	ATCGCCCGCGCAGAGAGAGCGGACGAGACGCGGTCATCGGCGGATCTGTCGACGACTG	885
QY	1252	CACGCCATGACATACGATACCGTACCGTGGTGTAGCTTCCACCGCGGCTGGTGCACAC	1311
DB	886	CAGTGTACCGCGAGGCGCGGTGCTCGCGTTCGGGACATGATGCGCAAGCGGATCACC	945
QY	1312	GTCAGAGGGCGGTACCGGCTTCCACATGGGCGCCCTGAACCCCGCAAGCTGGACGAG	1371
DB	946	ATGCGCGCGCACTCA---TGCAGAGCGCGCGGACATGATGACCTGTTCGAGCACTTCGC	1002

QY 1372 GCTGACGCGGACGCGCTGCGCCGACCGCTGCGCGCTGCGCGGAGG 1417
DB 1003 GCGCTCGCCGACGCGCTGCGCTGTACACCGCGCGGACTACGCGG 1048

RESULT 15

US-09-988-462-6
; Sequence 6, Application US/09988462
; Publication No. US20030046726A1
; GENERAL INFORMATION:

APPLICANT: Kozziel, Michael G.
Desai, Nalin M.
Lewis, Kelly S.
Kramer, Vance C.
Warren, Gregory W.
Evola, Stephen V.
Crossland, Lyle D.
Wright, Martha S.
Merlin, Ellis J.
Laudis, Karen L.

TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
INSECTICIDAL ACTIVITY IN MAIZE

NUMBER OF SEQUENCES: 94

CORRESPONDENCE ADDRESS:

ADDRESSEE: Syngenta Biotechnology, Inc.

STREET: 3054 Cornwallis Road

CITY: Research Triangle Park

STATE: NC

COUNTRY: USA

ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/988,462

FILING DATE: 20-NO. US20030046726A1-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/547,422

FILING DATE: 11-APR-2000

APPLICATION NUMBER: US 08/459,504

FILING DATE: 02-JUN-1995

APPLICATION NUMBER: US 07/951,715

FILING DATE: 25-SEP-1992

APPLICATION NUMBER: US 07/772,027

FILING DATE: 04-OCT-1991

ATTORNEY/AGENT INFORMATION:

NAME: Meigs, J. Timothy

REGISTRATION NUMBER: 38,241

REFERENCE/DOCKET NUMBER: S-188051

TELECOMMUNICATION INFORMATION:

TELEPHONE: (919)541-8587

TELEFAX: (919)541-8689

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 3624 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

DESCRIPTION: /desc = "Synthetic DNA"

HYPOTHETICAL: NO

FEATURE:

NAME/KEY: CDS

LOCATION: 1..3621

OTHER INFORMATION: /product= "Full-length, maize

optimized cryiB"

/note= "Disclosed in Figure 6."

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-988-462-6

Query Match 4.4%; Score 70; DB 9; Length 3624;
Best Local Similarity 45.4%; Pred. No. 7.7e-08;
Matches 293; Conservative 0; Mismatches 350; Indels 3; Gaps 1;

```
QY 926 AGGAGGCCCTGCAGCGGAGCTGGCGCTGCTGTGGACCCACCGCCCCCTGTTCGCCT 985
    ||| ||||| ||||| ||| ||| ||||| ||| |||
Db 2825 ACGACCGCTGCAGCGGAGACCAACATCGGCATGATCCACGCGCGGACAAAGCTGTGC 2884
    ||| ||||| ||||| ||| ||| ||||| ||| |||
QY 986 TCATCGCGCGCTGAGGAGCAGAGGCTGTGGACATCATCTGCGCGCCCTGCCAAGA 1045
    ||| ||||| ||||| ||| ||| ||||| ||| |||
Db 2885 ACCGCATCCGGAGGCCCTACTGAGCGAGCTGCCGTGATCCCGCGGTGACGCCGAGA 2944
    ||| ||||| ||||| ||| ||| ||||| ||| |||
QY 1046 TCCTGGCCACCCCAAGGTGAGATGCGCATCTCTGGGTACCGGAAGCCCGCTTACGAGA 1105
    ||| ||||| ||||| ||| ||| ||||| ||| |||
Db 2945 TCTTCGAGGAGCTGGAGGGCACATCATCACCGGCATCAGCCTGTACGACGCCGCAACG 3004
    ||| ||||| ||||| ||| ||| ||||| ||| |||
QY 1106 AGCTGGTGAACGCCATCGGCACCAAGTACAGAGGCCGCGCAAGGGCTGTCAAGTCT 1165
    ||| ||||| ||||| ||| ||| ||||| ||| |||
Db 3005 TGGTGAAGAACGGCGACTTCAACAACGCTGACTGCTGGAACGTGAAGGGCCACGTGG 3064
    ||| ||||| ||||| ||| ||| ||||| ||| |||
QY 1166 CGCGCGCCCTGGCGCACATGCTCACCGCGCGCGACTTCTGCTGGTGCCTCGCGCT 1225
    ||| ||||| ||||| ||| ||| ||||| ||| |||
Db 3065 ACCTGCACAGACCCACACCGCAGCGACTTGGTGATCCCGAGTGGGAGGCCGAGTGA 3124
    ||| ||||| ||||| ||| ||| ||||| ||| |||
QY 1226 TCGAGCCCTGGGCTGTATCCAGTGTGACG---CCATGCACTACGGTACCGTGCCTGG 1282
    ||| ||||| ||||| ||| ||| ||||| ||| |||
Db 3125 GCCAGCGGTGGCGGTGTGCCCGGCTGCGGCTACATCTGCGCGGTACCGCCTACAAG 3184
    ||| ||||| ||||| ||| ||| ||||| ||| |||
QY 1283 TAGCCTCCACCGCGGCTGTGTCGACACCGTCAAGGAGGCGTCAACGGCTTCCACATGG 1342
    ||| ||||| ||||| ||| ||| ||||| ||| |||
Db 3185 AGGGCTACGGGAGGGCTGCGTGACCATCCAGAGATCGAACAACACCGCAGCTGA 3244
    ||| ||||| ||||| ||| ||| ||||| ||| |||
QY 1343 GCGCCCTGAACCCCGCACAAAGCTGGACGAGGCTGACGCGGAGCCCTGGCGCGCACCGTGC 1402
    ||| ||||| ||||| ||| ||| ||||| ||| |||
Db 3245 AGTTCAAGAACCGGAGGAGGAGGTGTACCCACCGACACCGGCACCTGCCACGACT 3304
    ||| ||||| ||||| ||| ||| ||||| ||| |||
QY 1403 GCCGTGCCAGGAGGTGTTCGGGCGGCGCGCTACCCGAGATGTGGCCAACTGCATCA 1462
    ||| ||||| ||||| ||| ||| ||||| ||| |||
Db 3305 ACACCGCCACCAAGGGCACCGCGGCTGCGCGGAGCCCTGCAACAGCGCAACGCCGCT 3364
    ||| ||||| ||||| ||| ||| ||||| ||| |||
QY 1463 GCCAGGACCTGCTCTGTTCCAGCCCGGCCACAGTGGGAGGGCCTGCTGGAGGAGTGG 1522
    ||| ||||| ||||| ||| ||| ||||| ||| |||
Db 3365 ACGAGGAGCCTACGAGGTGACACCCCGCCAGCGTGAATCTACAAAGCCCACTACGAGG 3424
    ||| ||||| ||||| ||| ||| ||||| ||| |||
QY 1523 TGTACGGAAGGGCGCGCTGGCCACCGCCAAAGAGGAGATCAA 1568
    ||| ||||| ||||| ||| ||| ||||| ||| |||
Db 3425 AGGAGACCTACACCGACCTGGCGCGGACACCCACTGGGAGTACGA 3470
    ||| ||||| ||||| ||| ||| ||||| ||| |||
```

Search completed: June 6, 2003, 22:50:33
Job time : 161.611 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: June 4, 2003, 13:49:45 ; Search time 28.1631 Seconds
(without alignments)
2512.365 Million cell updates/sec

Title: US-09-980-771A-9
Perfect score: 2769
Sequence: 1 ALDIVVAEAPVPSKTGGL.....GGVATAKKEIKVPAEKIP 531

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1440	52.0	606	AA25476	Granule-bound star
2	1397.5	50.5	609	AAW56485	Oryza sativa starch
3	1390	50.2	527	AA25476	Arabidopsis thaliana
4	1390	50.2	563	AA25476	Arabidopsis thaliana
5	1390	50.2	610	AA25476	Arabidopsis thaliana
6	1390	50.2	610	AA25476	Arabidopsis thaliana
7	1353.5	48.9	533	AAW56484	Herbicide-activated
8	1344.5	48.6	637	AAW56484	Herbicide-activated
9	1331.5	48.1	502	AA25476	Protein encoded by
10	904	32.6	534	AA25476	Central fragment of

11	898	32.4	792	23	ABB92160	Herbicide-activated
12	891.5	32.2	647	20	AAW09004	Wheat starch solub
13	882.5	31.9	649	19	AAW38218	Maize starch synth
14	878	31.7	671	19	AAW23937	Maize starch synth
15	878	31.7	756	21	AAW50818	Wheat soluble star
16	873.5	31.5	626	15	AAW51231	Herbicide-activated
17	870.5	31.4	652	23	ABB93595	Soluble rice starch
18	861.5	31.1	539	19	AAW56491	Herbicide-activated
19	861.5	31.1	583	19	AAW56488	Herbicide-activated
20	852.5	30.8	583	19	AAW70894	Zea mays starch sy
21	847	30.6	802	23	AAW99845	Maize starch solub
22	847	30.6	813	23	AAU99844	Modified barley st
23	846	30.6	812	23	AAU99847	Barley cultivar Mo
24	845	30.5	804	19	AAW70892	Barley line MK6827
25	844	30.5	798	21	AAW37566	Maize starch solub
26	842	30.4	799	21	AAW37567	Wheat starch synth
27	842	30.4	799	21	AAW37597	Wheat starch synth
28	841	30.4	799	21	AAW37597	Wheat starch synth
29	837	30.2	698	19	AAW56487	Wheat granule-boun
30	825.5	29.8	812	23	AAU99846	Zea mays soluble s
31	825	29.8	597	21	AAW37568	Barley line 292 st
32	767	27.7	466	23	AAW51865	Wheat starch synth
33	726	26.2	459	17	AAW99540	Rice starch synth
34	701.5	25.3	669	19	AAW70893	Soluble starch syn
35	701.5	25.3	669	19	AAW56486	Maize starch solub
36	669	24.2	483	23	ABB54012	Zea mays soluble s
37	647.5	23.4	476	23	ABP27592	Lactococcus lactis
38	593	21.4	238	22	AAW31175	Streptococcus poly
39	580	20.9	477	13	AAW25462	A granule bound st
40	580	20.9	477	13	AAW25462	giga. Escherichia
41	568.5	20.5	677	19	AAW53890	E. coli glycogen s
42	568.5	20.5	1230	18	AAW99539	Soluble starch syn
43	568.5	20.5	1230	21	AAW17785	Potato tuber solub
44	561.5	20.3	1025	23	ABB90967	Potato starch synt
45	546	19.7	495	19	AAW70885	Herbicide-activated

ALIGNMENTS

RESULT 1
RAR25476
ID AAR25476 standard; Protein; 606 AA.
XX
AC AAR25476;
XX
XX
15-JAN-1993 (first entry)
XX
DE Granule-bound starch synthase of potato.
XX
KW GBSS gene; amylopectin; amylose production;
XX inhibition; promoter; antisense construct.
XX
OS Solanum tuberosum.
XX
XX
PN W09211376-A.
XX
XX
PD 09-JUL-1992.
XX
XX
PF 20-DEC-1991; 91WO-SF00892.
XX
XX
PR 21-DEC-1990; 90SE-0004096.
XX
XX
(AMYL-) AMYLOGENE HB.
XX
XX
PI Hofvander P, Persson PT, Tallberg A, Wikstrom O;
XX
XX
DR WPI; 1992-250096/30.
XX
XX
XX N-PSDB; AAQ26404.
XX
XX
PT Modifying potatoes to form amylopectin starch - using an
XX anti-sense construct to inhibit granule-bound starch synthase

Claim 6; Page 28-33; 46pp; English.

PS A genomic library in EMBL3 was prepared using leaves of the potato
 XX Blintje. The library was screened with cDNA clones for the 5' and 3'
 CC ends of the GBSS gene. A full-length clone of potato GBSS gene was
 CC identified (wx311) and isolated from the genomic library. The gene
 CC contained 12 introns. The amino acid sequence was deduced from it.
 CC Three fragments of the full-length sequence (5'-end, middle and
 CC 3'-end sequences) were characterised as suitable for use in novel
 CC antisense constructs to suppress amylose formation in potatoes.
 CC See also Q28400-3.

XX Sequence 606 AA;

Query Match 52.0%; Score 1440; DB 13; Length 606;
 Best Local Similarity 53.5%; Pred. No. 9.1e-116;
 Matches 292; Conservative 67; Mismatches 145; Indels 42; Gaps 10;
 QY 2 LDIWVAEAPWPKTGGGLDVTGGLPIELVKRHRVMTIAPRYDQYADAWDTSVVYDI- 60
 Db 81 MNLIFVTEGVPWKSITGELDVLGGLPPALAAAGHRVMTISPRYDQYKDAWDTGVAVEK 140
 QY 61 MG---EKVRPHSIKKGVRHVDHPFLAKVWGKTSKLYGPRSGADYLDNHRKRFALFC 117
 Db 141 VGSISIEIVRFHCYKRGVDRFVDHPMELEKVGKTSKLYGPRSGADYLDNHRKRFALFC 200
 QY 118 KAATEAARVLPE-----GP-GEDCVFVANDHWSALVPVLLKDEYQPKQFTKAKSVLA 169
 Db 201 QAALAPKVLNLSNYSFSGYEDVLFANDHWTALIPCYLKSMSYSGIYLNKAVAF 260
 QY 170 IHNIAFOGRWEEAFKDKLPQAFDKLAFSDGKAVYVTEATPMEDEKPLTGKTYKKI 229
 Db 261 IHNIAFOGRWEEAFKDKLPQAFDKLAFSDGKAVYVTEATPMEDEKPLTGKTYKKI 302
 QY 230 NWLKGIIADKLVTVSPNATEIAADAAGVVELDTVIRAKIGIIVNGMDIEENPKTD 289
 Db 303 NWMKAGILEADRLTVSPNATEIAADAAGVVELDTVIRAKIGIIVNGMDIEENPKTD 362
 QY 290 KFLSAPYDQNSVYAGKAAKAEALQALQELGPDVPTAPLFAFVIRGEEQKGVDIILALPKI 349
 Db 363 KYTDVKYDITVMDAKPILKEALQAAVGLPVDKXIPILGIFGRLEEKGSDILVAAIHKF 422
 QY 350 LATPKVQIATLGTKAAEKLVAIGTKYKRAKGVKFSAPLAHMLTAGADFMVPSRF 409
 Db 423 IGL-DVQITVLGIGTKKEFEQIEQLEVLNPNKAKGVAKFNVPPLAHMITAGADFMVPSRF 481
 QY 410 EPCGLIOLHAMHYGTVPVASTGGVLDTVKGVTFGFMGALNP--DKLDEADADALAAATV 467
 Db 482 EPCGLIOLHAMHYGTVPVASTGGVLDTVKGVTFGFMGALNP--DKLDEADADALAAATV 541
 QY 468 RRASEVFAAGGRYPPEMVANCISQDLSWSKPAQKWEGLL-----EEVYKGGVATAK 519
 Db 542 ARALAVYGTTLAFAEMIKNCMSEELSWKEPAKKWETLLGLGASGSEPGVEGE-ETAPLAK 600
 QY 520 EEIKVP 525
 Db 601 ENVATP 606

RESULT 2

AAW56485

ID AAW56485 standard; Protein: 609 AA.

XX AAW56485;

AC

XX 11-SEP-1998 (first entry)

DT

XX Oryza sativa starch (bacterial glycoen) synthase.

DE SER; starch-encapsulating region; fusion vector;

XX starch synthase; bacterial glycoen.

KW

XX Oryza sativa.

OS

XX WO9814601-A1.

PN

XX 09-APR-1998.

PD

XX 30-SEP-1997; 97WO-US17555.

XX

XX 30-SEP-1996; 96US-0026855.

XX

XX (EXSE-) EXSEED GENETICS LLC.

XX

XX Guan H, Keeling P;

XX

XX WPI; 1998-240100/21.

XX

XX N-PSDB; AAV29753.

XX

XX Hybrid polypeptide comprising starch-encapsulating region and

XX protein, -useful for, e.g. producing protein(s) resistant to

XX degradation by stomach acids

XX

XX Example 2; Page 32; 156pp; English.

XX

XX The sequence is that of starch (bacterial glycoen) which is

XX encoded by the waxy gene. It can be used in the production of a hybrid

XX polypeptide comprising a starch-encapsulating region (SER) fused

XX to a payload protein. The hybrid polypeptide can be used to make

XX modified starches comprising the payload protein, selected from,

XX e.g. hormones, growth factors, antibodies, enzymes, dyes,

XX immunoglobulins, etc. The modified starch can also be used

XX to provide grain feeds enriched in amino acids. By encapsulating

XX the payload protein in starch, it is more resistant to

XX degradation by stomach acids.

XX

XX Sequence 609 AA;

QY Query Match 50.5%; Score 1397.5; DB 19; Length 609;

Best Local Similarity 51.4%; Pred. No. 4.4e-112;

Matches 281; Conservative 76; Mismatches 147; Indels 43; Gaps 9;

QY 2 LDIWVAEAPWPKTGGGLDVTGGLPIELVKRHRVMTIAPRYDQYADAWDTSVVYDI- 60

Db 83 MNVVFVGAEMAPWPKTGGGLDVTGGLPPAMAANGHRVMTISPRYDQYKDAWDTSVVAETK 142

QY 61 ---MGEKVRPHSIKKGVRHVDHPFLAKVWGKTSKLYGPRSGADYLDNHRKRFALFC 117

Db 143 VADRYERVRFFHCYKRGVDRFVDHPSELEKVGKTSKLYGPRSGADYLDNHRKRFALFC 202

QY 118 KAATEAARVL-----PFGP---GEDCVFVANDHWSALVPVLLKDEYQPKQFTKAKSVLA 169

Db 203 QAALAPKVLNLSNYSFSGYEDVLFANDHWTALIPCYLKSMSYSGIYLNKAVAF 262

QY 170 IHNIAFOGRWEEAFKDKLPQAFDKLAFSDGKAVYVTEATPMEDEKPLTGKTYKKI 229

Db 263 IHNIAFOGRWEEAFKDKLPQAFDKLAFSDGKAVYVTEATPMEDEKPLTGKTYKKI 305

QY 230 NWLKGIIADKLVTVSPNATEIAADAAGVVELDTVIRAKIGIIVNGMDIEENPKTD 289

Db 306 NWMKAGILEADRLTVSPNATEIAADAAGVVELDTVIRAKIGIIVNGMDIEENPKTD 365

QY 290 KFLSAPYDQNSVYAGKAAKAEALQALQELGPDVPTAPLFAFVIRGEEQKGVDIILALPKI 349

Db 366 KYITAKYDITVMDAKPILKEALQAAVGLPVDKXIPILGIFGRLEEKGSDILVAAIHKF 425

QY 350 LATPKVQIATLGTKAAEKLVAIGTKYKRAKGVKFSAPLAHMLTAGADFMVPSRF 409

Db 426 M-OEDVQIVLLGIGTKKEFEQIEQLEVLNPNKAKGVAKFNVPPLAHMITAGADFMVPSRF 484

QY 410 EPCGLIOLHAMHYGTVPVASTGGVLDTVKGVTFGFMGALNP--DKLDEADADALAAATV 467

Db 485 EPCGLIOLHAMHYGTVPVASTGGVLDTVKGVTFGFMGALNP--DKLDEADADALAAATV 544

QY 468 RRASEVFAAGGRYPPEMVANCISQDLSWSKPAQKWEGLLEEVYKGGVATAK 518

```

Db      545 KRAIKVGTAYEEMVRNMCNODLSWKGFKNWNL--LGLGVAGSAPCIEGDEIAPLA 602
QY      519 KEEIKVP 525
        |||::|
Db      603 KENVAAP 609

RESULT 3
AAG04668
ID AAG04668 standard; Protein; 527 AA.
XX
AC AAG04668;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 782.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0143977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.

```


PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0132048.
PR 21-JUL-1999; 99US-0132407.
PR 22-JUL-1999; 99US-0132484.
PR 22-JUL-1999; 99US-0132485.
PR 22-JUL-1999; 99US-0132486.
PR 22-JUL-1999; 99US-0132487.
PR 22-JUL-1999; 99US-0132488.
PR 23-JUL-1999; 99US-0132489.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148441.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

QY 287 KTDKFLSAPYDQNSVYAGRAAAKEALQAEGLGVDPTAPLFAFGRLEEQKGVDDIILAAAL 346
DB 364 STDKYIDIKYDITVTDAKPLIKEALQAAVGLPVDPRVPVIGFGRLEEQKGVDDIILAAAL 423
QY 347 PKILATPKVOJAILGTGKAAYEKLVAIGYKGRAGVVKFSAPLAHMTAGADFMLVP 406
DB 424 SKFMGL-NVQMVILGTGKKAAYEKLVAIGYKGRAGVVKFSAPLAHMTAGADFMLVP 482
QY 407 SREPECGLIQHAMHYGTVPVASTGGLVDTVKGVTFHMGALNP--DKLDEADADALA 464
DB 483 SREPECGLIQHAMRYGTVPVASTGGLVDTVKGVTFHMGALNP--DKLDEADADALA 542
QY 465 ATVRRASEVAGGRYPENVANCISQDLSWSKPAQKWEGLLEEVYKGVGATAKKEIKV 524
DB 543 KAVTRAVAVYGTSAQEMVKNQMDQDFSWKGPRLWEKVLVSLNVA-GSEAGTEGEEI-A 600
QY 525 PVAEK 529
DB 601 PLAKE 605

RESULT 6
ABB91229
ID ABB91229 standard; Protein; 610 AA.
XX ABB91229;
XX
XX 31-MAY-2002 (first entry)
XX
XX Herbicidally active polypeptide SEQ ID NO 440.
XX
XX Herbicidal; plant; agriculture; herbicide.
XX
XX Arabidopsis thaliana.
XX
XX WO200210210-A2.
XX
XX 07-FEB-2002.
XX
XX 28-AUG-2001; 2001WO-EP09892.
XX
XX 28-AUG-2001; 2001WO-EP09892.
XX
XX (FARB) HAYER AG.
XX
XX Tietjen K, Weidler M;
XX
XX WPI; 2002-269010/31.
XX
XX Identifying plant target proteins for herbicidally active compounds,
XX comprising aligning and comparing nucleic acid or amino acid sequences
XX from plant with nucleic acid or amino acid sequences from non-plant
XX organisms -
XX
XX Claim 5; SEQ ID NO 440; 261pp + Sequence Listing; English.
XX
XX The invention relates to identifying target proteins
XX (ABB90790-ABB94016) for herbicidally active compounds, comprising
XX aligning and comparing nucleic acid or amino acid sequences from plant
XX with nucleic acid or amino acid sequences from non-plant organisms using
XX suitable search parameters, where plant sequences having an E-value
XX greater by a factor of 3 than the E-value of most similar non-plant
XX sequences are selected. The polypeptides or nucleic acids encoding them
XX are useful for identifying modulators. The identified modulators are
XX useful as herbicides.
XX
XX Sequence 610 AA;
XX

Query Match 50.2%; Score 1390; DB 23; Length 610;
Best Local Similarity 52.7%; Pred. No. 2e-111;
Matches 287; Conservative 72; Mismatches 146; Indels 40; Gaps 12;
QY 2 IDIVVAEAVFWSKGTGLGVDVIGLPIELVKRHRVMTIAPRYDQYADAWTTSVVVDI- 60

DB 84 MSVIFIGAEVGPWSKTGGLGVDLPPALAAHRVMTICPRVDQYKADMTCCVVOIK 143
QY 61 MGEX---VRYFHSIKKGVHRVWIDHPWFLAKVWKGKGLYKPRSGADYLDNHRKRFALFC 117
DB 144 VGDKVENVRPHCYKRGVDRVFDHPFLAKVWKGKGLYKPRSGADYLDNHRKRFALFC 203
QY 118 KAAIEAARVLPF-----GP-GEDCVFVANDHNSALVPLVLLKDEYQKGFQTRAKSVLA 169
DB 204 QAALAEQVNLNLSKYSFSGPYGEDVVFVANDHNTALLPCYLKSMYQSRGVYMAKVVFC 263
QY 170 IHNTAFQGRMWEAEKDKLP---QAADFKLAFSDGYAKVYTEATPMEDEKPLTGKTY 226
DB 264 IHNTAYQGRFAFDYSLNLPISPKSSFD---FMDGYEK-----PVKG--- 303
QY 227 KKNLWKGIIAADKLVTPSPNATEIAADAAGGVVELDTVIRAKGIEGIVNGMDIEEWN 286
DB 304 RKINWKAAILLEAHRVLTSPYQAELISGVDRGVLELHKYLRMTKTVSGIINGMDVQEWNP 363
QY 287 KTDKFLSAPYDQNSVYAGRAAAKEALQAEGLGVDPTAPLFAFGRLEEQKGVDDIILAAAL 346
DB 364 STDKYIDIKYDITVTDAKPLIKEALQAAVGLPVDPRVPVIGFGRLEEQKGVDDIILAAAL 423
QY 347 PKILATPKVOJAILGTGKAAYEKLVAIGYKGRAGVVKFSAPLAHMTAGADFMLVP 406
DB 424 SKFMGL-NVQMVILGTGKKAAYEKLVAIGYKGRAGVVKFSAPLAHMTAGADFMLVP 482
QY 407 SREPECGLIQHAMHYGTVPVASTGGLVDTVKGVTFHMGALNP--DKLDEADADALA 464
DB 483 SREPECGLIQHAMRYGTVPVASTGGLVDTVKGVTFHMGALNP--DKLDEADADALA 542
QY 465 ATVRRASEVAGGRYPENVANCISQDLSWSKPAQKWEGLLEEVYKGVGATAKKEIKV 524
DB 543 KAVTRAVAVYGTSAQEMVKNQMDQDFSWKGPRLWEKVLVSLNVA-GSEAGTEGEEI-A 600
QY 525 PVAEK 529
DB 601 PLAKE 605

RESULT 7
AAW56484
ID AAW56484 standard; Protein; 533 AA.
XX AAW56484;
XX
XX 11-SEP-1998 (first entry)
XX
XX Zea mays waxy gene glucosyl transferase.
XX
XX SER; starch-encapsulating region; fusion vector.
XX
XX Zea mays.
XX
XX WO9814601-A1.
XX
XX 09-APR-1998.
XX
XX 30-SEP-1997; 97WO-US17555.
XX
XX 30-SEP-1996; 96US-0026855.
XX
XX (EXSE-) EXSEED GENETICS LLC.
XX
XX Guan H, Keeling P;
XX
XX WPI; 1998-240100/21.
XX
XX N-PSDB; AAV29752.
XX
XX Hybrid polypeptide comprising starch-encapsulating region and
XX protein - useful for, e.g. producing protein(s) resistant to
XX degradation by stomach acids

Example 2; Page 28; 156pp; English.

The sequence is that of maize glucosyl transferase which is encoded by the waxy gene. It can be used in the production of a hybrid polypeptide comprising a starch-encapsulating region (SER) fused to a payload protein. The hybrid polypeptide can be used to make modified starches comprising the payload protein, selected from, e.g. hormones, growth factors, antibodies, enzymes, dyes, immunoglobulins, etc. The modified starch can also be used to provide grain feeds enriched in amino acids. By encapsulating the payload protein in starch, it is more resistant to degradation by stomach acids.

Sequence 533 AA;

```

Query Match          48.9%; Score 1353.5; DB 19; Length 533;
Best Local Similarity 50.4%; Pred. No. 2.3e-108;
Matches 275; Conservative 79; Mismatches 153; Indels 39; Gaps 11;

QY 2 LDIVMAAEVAPWSKGTGLGSDVTGGLPLEIKRGRHVMNTIAPRYDOYADAMDTSVVVDI- 60
DB ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
5 MNVVFVGAEMAPWSKGTGLGSDVGLGLPPAMAANGHRVMVSPRYDQYKDAWDTSVVSEIK 64
QY 61 MG---EKVRVPHSITKKGVRHVRWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKRFALFC 117
DB ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
65 MGDGYETVRPFHCYKRGVDRVFDVDFLPLERVWCKTEKIYGPVAGTDYRDNQLRFSLLC 124
QY 118 KAATEAARVL-----PF--GP-GEDCVFVANDHWSALVPVLLKDEYQPKGFTKAKSVLA 169
DB ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
125 QAALAPRILSLNNPYPFSGPYGDEVDVFCNDWHTGTLSCYKSNYQSHGYRDAKTAF 184
QY 170 IHNIAFOGRMWEAEFKDTKLPOAAFDKLAISDGAKVYVTEATPWEDEKPLTCKTYKKI 229
DB ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
185 IHNISYOGREAFSDYPELNLPERKSSFDIDGVEK-----PVEG---RKI 227
QY 230 NLWKGGLIADKLVTVSPNYATEITAADAAGVELDTVIRAKGIEGIVNGMDIEWNP 289
DB ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
228 NWMKAGILEADRLVTVSPYAEELISGIARCELDNIMRLTGINGVNGMDVSEWDSRD 287
QY 290 KFLSAPYDONSIVYAGKAAKEALQAEGLPVDPTAPLFAFIRLEEQKGVDDIILAALPKI 349
DB ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
288 KYIAVKYDVSTAVAEKALNEALQAEVGLPVDRIPLVAFIRLEEQKGVDPVMAAALPQL 347
QY 350 L-ATPKVQIAILGTGKAAYEKLVNAICTKYKGRAGVVKFSAPLAHMLTAGADFMVPSR 408
DB ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
348 MEMVEDQVILLGTGKKKEFERMLMSAEKFPKGVRAVVKFNAAALAHMAGADVLA 407
QY 409 FEPGGLIOLHAMHYGVVVPVASTGGLVDYTKEGVTGFHMGALNPD--KLDEADADALAAT 466
DB ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
408 FEPGGLIOLQGMRYGTPCASCAGGLVDYTIIEGTGFGHMGRLSYDCNVNVEPADVKVAT 467
QY 467 VRASEVFAAGRYPEMVANCISODLSWSKPAQWEGLEEE--VYVKGGV-----ATAKK 519
DB ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
468 LQRAIKVGTGTPATEEMVRNMIODLSWKGPANWENLLSLGAVGAGGEPGVEGETIAPLAK 527
QY 520 ERIKVP 525
DB ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
528 ENVAAP 533

```

RESULT 8

AA116604

ID AAY16604 standard; Protein; 637 AA.

AC AAY16604;

XX AAY16604;

DT 23-AUG-1999 (first entry)

XX Protein encoded by the DNA sequence of the maize waxy gene.

DE Non-glycogen-like polysaccharide production; fermentation; waxy gene.

XX starch synthesis enzyme; glycogen-synthesis enzyme; glycogen synthesis;

KW non-starch branching gene; amylopectin; amylose; plant-like starch.

KW

XX OS
XX Zea mays.
XX WO9844780-A1.
XX 15-OCT-1998.
XX 03-APR-1998; 98WO-US06660.
XX 04-APR-1997; 97US-0042939.
XX (EXSE-) EXSEED GENETICS LLC.
XX Guan H, Keeling PL;
XX WPI; 1998-568285/48.
XX N-PSDB; AAX60319.
XX Producing non-glycogen-like polysaccharides in bacteria, fungi or
XX plants - transformed with genes for enzymes involved in starch or
XX glycogen synthesis allows fermentative production of starches with
XX engineered properties
XX Disclosure; Fig 49; 150pp; English.
XX The specification describes a method for the production of
XX non-glycogen-like polysaccharides in a host. The method comprises
XX transforming a host, suitable for fermentation, with genes encoding
XX starch- or glycogen-synthesis enzymes, and fermenting the transformants.
XX The specification also describes hosts transformed with a gene active
XX in glycogen synthesis and at least one non-starch branching gene,
XX involved in production of amylopectin or amylose in its original host.
XX The method is used to produce plant-like starches by fermentation and
XX new starches in plants. These starches are useful for all food and
XX non-food applications of starch. The present sequence is used in
XX the course of the invention.
XX Sequence 637 AA;

Query Match

Best Local Similarity 48.6%; Score 1344.5; DB 19; Length 637;

Matches 273; Conservative 79; Mismatches 155; Indels 39; Gaps 11;

```

QY 2 LDIVMAAEVAPWSKGTGLGSDVTGGLPLEIKRGRHVMNTIAPRYDOYADAMDTSVVVDI- 60
DB ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
77 MNVVFVGAEMAPWSKGTGLGSDVGLGLPPAMAANGHRVMVSPRYDQYKDAWDTSVVSEIK 136
QY 61 MG---EKVRVPHSITKKGVRHVRWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKRFALFC 117
DB ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
137 MGDGYETVRPFHCYKRGVDRVFDVDFLPLERVWCKTEKIYGPVAGTDYRDNQLRFSLLC 196
QY 118 KAATEAARVL-----PF--GP-GEDCVFVANDHWSALVPVLLKDEYQPKGFTKAKSVLA 169
DB ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
197 QGALEAPRILSLNNPYPFSGPYGDEVDVFCNDWHTGTLSCYKSNYQSHGYRDAKTAF 256
QY 170 IHNIAFOGRMWEAEFKDTKLPOAAFDKLAISDGAKVYVTEATPWEDEKPLTCKTYKKI 229
DB ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
257 IHNISYOGREAFSDYPELNLPERKSSFDIDGVEK-----PVEG---RKI 299
QY 230 NLWKGGLIADKLVTVSPNYATEITAADAAGVELDTVIRAKGIEGIVNGMDIEWNP 289
DB ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
300 NWMKAGILEADRLVTVSPYAEELISGIARCELDNIMRLTGINGVNGMDVSEWDSRD 359
QY 290 KFLSAPYDONSIVYAGKAAKEALQAEGLPVDPTAPLFAFIRLEEQKGVDDIILAALPKI 349
DB ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
360 KYIAVKYDVSTAVAEKALNEALQAEVGLPVDRIPLVAFIRLEEQKGVDPVMAAALPQL 419
QY 350 L-ATPKVQIAILGTGKAAYEKLVNAICTKYKGRAGVVKFSAPLAHMLTAGADFMVPSR 408
DB ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
420 MEMVEDQVILLGTGKKKEFERMLMSAEKFPKGVRAVVKFNAAALAHMAGADVLA 479
QY 409 FEPGGLIOLHAMHYGVVVPVASTGGLVDYTKEGVTGFHMGALNPD--KLDEADADALAAT 466
DB ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```


Query Match 32.6%; Score 904; DB 21; Length 534;
 Best Local Similarity 32.5%; Pred. No. 1.8e-69;
 Matches 213; Conservative 48; Mismatches 103; Indels 292; Gaps 9;

QY 7 VAAEVAPWKTGGLGDTGGTGLPIELVKGHRVMTIAPRYDOYADAWDTSVVVDIMGEKVR 66
 DB 2 VGEAMAPWKTGGLGDLGLGPPMAANGHRVMTISPRDYQYKDAWDT----- 49
 QY 67 YFHSIKKGVRWIDHPFLAKWGTGSKLYGPRSGADYLDNHRKFALEFCRAALEAARV 126
 DB 50 -----LEKVRGKTKKIYDPDAGTDYEDNQRFSLCQAALVEPRI 90
 QY 127 I-----PF----- 129
 DB 91 LNLNDNPFYFCHNISIYGRSFDFDFAQLNLPDRKSSFDFTDGYDKPVEGRKISVSEI 150
 QY 130 -----GP-GECDVCFVANDHSALEVPVLLKDEYOPK 158
 DB 151 KVVDKYERVRYPHCYKRGVDRVVDHPCFSGPYGEDVFCVNDWHTGLLACLYLKSNYOSN 210
 QY 159 GQFTKAKSVLAHNTAFQGRWEEAFKDTKLPQAAFDKLAESDGYAKYVTEATPMEDEK 218
 DB 211 GIYRAAKVA----- 219
 QY 219 PPLTGKTYKIKNLKGGIILADKLVTSPNVALEIAADAAGGVLDTVIRAKGIGVNG 278
 DB 220 -----NWKAGILOADKLVITSPYTAELISGEARGCELDNIMRLTIGTITVNG 268
 QY 279 MDIEWNPKTDFLSAPYDQNSVYAGAAAK----- 309
 DB 269 MDVSEWDPKDFLAVNYDITITALEKALNEALEGKALNKIKEEDVOIVILGTGKKFE 328
 QY 310 -----EALQAEGLPVDPTAPLFAFGRLEEQKGVDTILA 344
 DB 329 RLLKSIEKFSKRVAVVRFNAPLAELALQAEVGLPVDKRVPLVAFIGRLEEQKGPVMA 388
 QY 345 ALPKILATPKVOJAILGTGAAEYKLVNAIGTKYKGRAGVVKFSAPLAHMLTAGADPML 404
 DB 389 SIPEI-----HOMMAGADVIA 404
 QY 405 VPSFEPCGLIQLHAMHYGTPVVAAGGLVDTKEGVTGPHMGALNPD--KLDEADADA 462
 DB 405 VTSFEPCGLIQLQMGYGTGPCACASTGGLVDTIVEGKTGPHMGLSYDCNVVPEADYK 464
 QY 463 LAATVRASEVFGAGRYPE-----WVANCISQDLSWSKPAOKWEGLLBEV 507
 DB 465 VVTLKRAVVGTPAYTHEAPLAMENVAAPMVKNCMIQDLSWKGPAKNWEDVLEL 520

RESULT 11
 ID ABB92160
 AC ABB92160;
 XX ABB92160;
 DT 31-MAY-2002 (first entry)
 DE Herbicidally active polypeptide SEQ ID NO 1371.
 XX Herbicidal; plant; agriculture; herbicide.
 KW Arabidopsis thaliana.
 OS Arabidopsis thaliana.
 PN W0200210210-A2.
 XX W0200210210-A2.
 PD 07-FEB-2002.
 XX 07-FEB-2002.
 PF 28-AUG-2001; 2001WO-EF09892.
 XX 28-AUG-2001; 2001WO-EF09892.
 PR 28-AUG-2001; 2001WO-EF09892.
 XX 28-AUG-2001; 2001WO-EF09892.
 PA (FARB) BAYER AG.

XX Tietjen K, Weidler M;
 XX WPI; 2002-269010/31.
 XX Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -
 XX
 XX Claim 5; SEQ ID NO 1371; 261pp + Sequence Listing; English.
 XX The invention relates to identifying target proteins
 CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.
 XX
 XX Sequence 792 AA;
 QY 2 LDIVVAAEVAWPWKTGGLGDTGGTGLPIELVKGHRVMTIAPRYDOYADAWDTSV--VVD 59
 DB 301 MNVILVAAEAFPFSGTGLGVDAGALPKSLARRGHRVMTIAPRYAEAKDLGVRKRYK 360
 QY 60 IMGE--KYRPHSIIKKGVRWIDHPFLAKWGTGSKLYGPRSGADYLDNHRKFALEFC 117
 DB 361 VAGQDMYVYHAFIDGVDFVDFIDSEF-----RHLSNNIY----GGNRKLDITLKRKVLFC 411
 QY 118 KAAIEAARVLP-----FGPEDCVFVANDHSALEVPVLLKDEYQPKGQFTAKSVLAHNL 172
 DB 412 KAAVEVPWVPCGVGCGYGDG--NLAFIANDWHTALLPVYLKAYYRDHGMKTYRSLVITHN 470
 QY 173 IAFGRWKEEAFKDTKLPQAAFDKLAESDGYAKYVTEATPMEDEKPLTKTYKIKINWL 232
 DB 471 IAHOGRGVDDFSYVDLPSHYLDSFKLYD-----PVGGEHF---NIF 509
 QY 233 KGGIITADKLVTSPNVALEIAADAAGGVLDTVIRAKG--IEGIVNGMDIEEWNPKYDK 290
 DB 510 AAGLKAADRVLTSHGYSWEVKT--LEGGWGLHNIINENDWKFGRGIVNGIDTQWNPEDT 568
 QY 291 FLSAP----YDQNSVYAGKAAKAEALQAEGLPVDPTAPLFAFGRLEEQKGVDTILAAL 346
 DB 569 YLHSDDTYNTYSLNHLHGKPCQAALQELGLPVRPDVPLIGFGRLDHKGVDLIAEAV 628
 QY 347 PKILATPKVOJAILGTGKAAEYKLVNAIGTKYKGRAGVVKFSAPLAHMLTAGADPMLVP 406
 DB 629 PWWW-SQDVQLVMLGTGRPDLEEVLRQMEHOYRDKRCWGVGFSVKTHRITAGADILMP 687
 QY 407 SRFEPCGLIQLHAMHYGTPVVAAGGLVDTKEGVTGPHMGALNPDKLDEADA 460
 DB 688 SRFEPCGLNQLYANNYGTIPVVAAGGLRDVQVQFDYPSETGLGWTEDSAGKLIHALG 747
 QY 461 DALAATVRASEVFGAGRYPEWVANCISQDLSWSKPAOKWEGLLBEVVY 509
 DB 748 NCL-LTYREYKESWEG-----LQRRGMTQDLSWDAEAEKVEEVLVAARY 790

RESULT 12
 ID AAY09004
 AC AAY09004;
 XX AAY09004;
 DT 05-JUL-1999 (first entry)
 XX

Db 202 HIRIPCGGEHEVTFEHEYRDSVDVFDHPSPY-----HRPGLNYGDKFGA-TGDNQFR 254
 QY 113 FALFCKAAEAPVLPFGP----GECVFNWHSALVPVLLKDEYQPKGQFTAKSVIA 169
 Db 255 YTLICAAACEAPLVLELGGYTYGQNCMEVNDWHSALVPVLLAARYGYGVYKDSRSLIV 314
 QY 170 IHNIAFGRMWEEAFKDKLPQAAFDKLAFGSDGYAKVYTEATPMEDEKPLTKYKIKI 229
 Db 315 IHNLAHOGVEPASTYPLDGLPPEWYGALEW-----VPEWARHLDKG-----EAV 361
 QY 230 NLWKGIIADKLVTPSPNYATEIAADAAGGVLDTVI--RAKGIEGVNMDIEENPK 287
 Db 362 NFLKGAVTADRIVTSYKGSWEVTT-AEGGQGLNELLSRKSVLNGVINGIDINDWPA 420
 QY 288 TDFELSAFYDQNSVYAGKAAKEALQAEGLPVDPTAPLFAFGRLEEQKGVDIILALP 347
 Db 421 TDKCIPCHYSVDDL-SGAKCKALQELGLPIRPDVLIGFIGRLDYQKGLDILQIIP 479
 QY 348 KILATPKVQIAILGTGKAAVEKLVNAIGTKYGRAGKGVVKSAPLAHMLTAGADFMVPS 407
 Db 480 DLM-REDVQFVNLGSGDPELEDWNRSTESIFKDKFGWVGFSVPVSHRITAGCDILLMS 538
 QY 408 RFPCGLIQLIHAMHYGTPVNVASTGGVLDTVK-----EGVTGFHMGALNPKLDEA 458
 Db 539 RFPCGLNOLYAMOGYTPVNVHATGGLROTVENFNPFENGEGQCTGWAFAPLTEN---- 594
 QY 459 DADALATVRRASEVAGGR--YPEMVANCISQDLSKSPAKWEGILLE 505
 Db 595 -----MLWTLTALSTYREHKSSEGLMKRMSKDFTWDAARQVEIQIF 639

RESULT 14

AAW23937
 ID AAW23937 standard; Protein; 671 AA.

AC AAW23937;

XX 21-MAY-1998 (first entry)
 XX

DE Wheat soluble starch synthase partial sequence.

XX Starch synthase; wheat; transgenic plant.

XX Triticum aestivum L. cv. Florida.

XX WO9745545-A1.

XX 04-DEC-1997.

XX 28-MAY-1997; 97WO-EF02793.

XX 11-SEP-1996; 96DE-1036917.

XX 29-MAY-1996; 96DE-1021588.

XX (AGRE) HOECHST-SCHERING AGREVO GMBH.

XX Block M, Loerz H, Luetticke S, Froberg C, Kossmann J;
 PI Walter L;

XX WPI; 1998-032652/03.

XX N-PSDB; AAV01527.

XX Nucleic acid encoding starch synthase enzymes from wheat - for
 PT transgenic plants that produce modified forms of starch, useful e.g.
 PT in foods, or for production of packaging materials and disposable
 PT goods

XX Claim 1; Page 47-50; 71pp; English.

XX This amino acid sequence comprises a near full-length sequence for
 CC a soluble starch synthase of summer wheat (cv. Florida). It was
 CC deduced from a cDNA clone (AAV01527) isolated from a 21-day caryopsis
 CC cDNA library. A granule-bound starch synthase (see AAW23938) has

CC also been identified. Isolated nucleic acids encoding these
 CC enzymes can be inserted into vectors for production of transgenic
 CC plants, particularly starch-producing plants, specifically wheat.
 CC Use of the isolated nucleic acids, or of antisense sequences, allows
 CC starch metabolism to be regulated in transgenic plants.
 CC Overexpression may result in improved crop yield, while modification
 CC of starch in planta may eliminate the need for subsequent
 CC chemical/physical modification. Plants with altered levels of the
 CC various isoforms of starch synthase will produce starch of different
 CC chain length, amylose/amylopectin ratio, degree of branching,
 CC phosphate content, gelatinisation behaviour, granule size and shape,
 CC viscosity etc. The starch produced by such plants is useful
 CC particularly in foods (especially bakery goods or pasta) or to
 CC produce packaging materials or disposable goods, as well as in any
 CC other known use of starch.

XX Sequence 671 AA;

Query Match 31.7%; Score 878; DB 19; Length 671;
 Best Local Similarity 39.9%; Pred. No. 4.5e-67;
 Matches 210; Conservative 78; Mismatches 164; Indels 74; Gaps 16;

QY 4 IVMVAAEAVAPSKTCGLGDVTGGPIELVKGHRVMTIAPRY-----DQYADAWDTSVV 57
 Db 56 IVFVTGEAAPYAKSGGLGDCGSLFIALAARHVRVMPVRYLNGSSDKNTAKALYTAH 115
 QY 58 VDIM---GEKVRYFHSIKGVHRVWIDHPWFLAKVNGKTSKLYGPRSGADYLDNHRF 113
 Db 116 IKIPCFGSHVETFEHYRDNVDVFDHP-----SYHRPGS-LYGNFGA-FGDNQFRY 168
 QY 114 ALFCKAAEAPVLPFGP---GEDCVFNWHSALVPVLLKDEYQPKGQFTAKSVIAI 170
 Db 169 TLLCYAAACEAPLVLELGGYTYGQNCMEVNDWHSALVPVLLAARYGYGVYRDSRSLIV 228
 QY 171 HNTAFQGRMWEAFKDKLPQAAFDKLAFGSDGYAKVYTEATPMEDEKPLTKYKIKIN 230
 Db 229 HNTAHOGVEPASTYPLDGLPPEWYGALEW-----VPEWARHLDKG-----EAVN 275
 QY 231 WLKGGIITADKLVTPSPNYATEIAADAAGGVLDTVI--RAKGIEGVNMDIEENPKT 288
 Db 276 FLKGAVTADRIVTSYKGSWEVTT-AEGGQGLNELLSRKSVLNGVINGIDINDWPT 334
 QY 289 DKFELSAFYDQNSVYAGKAAKEALQAEGLPVDPTAPLFAFGRLEEQKGVDIILALP 348
 Db 335 DKCLPHYSVDDL-SGAKCKALQELGLPVRDVLIGFIGRLDYQKGLDILKMAIPE 393
 QY 349 ILATPKVQIAILGTGKAAVEKLVNAIGTKYGRAGKGVVKSAPLAHMLTAGADFMVPSR 408
 Db 394 LM-REDVQFVNLGSGDPIFEGWRSTESSYKDKFRGVWVGFSVPVSHRITAGCDILLMPSR 452
 QY 409 FEPCGLIQLIHAMHYGTPVNVASTGGVLDTVK-----EGVTGFHMGALNPKLDEAD 459
 Db 453 FEPCGLNOLYAMOGYTPVNVHATGGLROTVENFNPFENGEGQCTGWAFAPLTEN---- 511
 QY 460 ADALATVRRASEVAGGRYPEMVANCISQDLSKSPAKWEGILLE 505
 Db 512 -----LRTAMSTFRE-----HKPS--WEGLMK 531

RESULT 15

AAV50818

ID AAV50818 standard; Protein; 756 AA.

AC AAV50818;

XX 18-FEB-2000 (first entry)

XX Wheat soluble starch synthase protein.
 XX

XX Soluble; starch synthase; wheat; transgenic plant; starch production;
 KW food; baking; pastry; packaging material; glucose; glucan; paper; pulp;
 KW adhesive; textile; building material; soil stabilizer; wetting agent;
 KW fertilizer; plant-protection; cosmetic; flocculant.

QY 349 ILATPKVOTAILGTGKAAEYKLYNAIGTKYGRAGVGVVTSAPLAHMLTAGADFMVPSR 408
Db 479 LM-REDVQFVYMLGSDPIFEQWKRSTESSYKDKFRGWGVSVVSHRITAGCDILLMPSR 537
QY 409 FEPCGLIOLHAMHYGIVPVVASTGGGLVDIVK-----EGVTGFHMGALNPKLDEAD 459
Db 538 FEPCGLNOLYAMOYGVVPHVHGGGLRDIVETNPFCAKGEETGWAFFSLTVDKMLWA- 596
QY 460 ADALAATVRRASEVFAAGGYPYEMVANCISQDLNSKSPAKQWEGILLE 505
Db 597 -----LRTAMSTFRE-----HKPS--WEGIMK 616

Search completed: June 4, 2003, 14:52:32
Job time : 32.1631 secs

XX Triticum aestivum.
XX DEL9820607-Al.
XX 11-NOV-1999.
XX 08-MAY-1998; 98DE-1020607.
XX 08-MAY-1998; 98DE-1020607.
XX (AGRE) HOECHST-SCHERING AGREVO GMBH.
XX Loerz H, Luetticke S, Block M;
XX WPI; 2000-024508/03.
XX N-PSDB; AAZ24487.
XX New enzyme with starch synthase activity, useful for producing starch
XX for foods and packaging materials -
XX Claim 1a: Page 19-21; 24pp; German.
XX This invention describes a novel protein (I) with the activity of wheat
XX starch synthase. Transgenic plants, specifically wheat, that contain (I)
XX are used for production of starch, used particularly in foods,
XX particularly baked and pastry goods and for making packaging materials or
XX disposable items. Starch may also be used as starting materials for
XX glucose or glucan components (e.g. for fermentation or further chemical
XX conversion); in paper and pulp production, as adhesives, in textiles,
XX in preparation of gypsum-based building materials, as soil stabilizer,
XX as wetting agent etc. in fertilizer and plant-protection compositions,
XX as binder (in pharmaceuticals, cosmetics, coal briquetting and casting
XX sand), as flocculant in soil or coal slurries, as rubber and leather
XX additives, and for production of synthetic polymers, e.g. polyurethane
XX films. Transgenic plants with increased/decreased production of (I)
XX produce starches with altered physical and/or chemical properties such as
XX amylose/amylopectin ratios, degree of branching, mean chain length,
XX phosphate content, gelatinization properties, gel- or film-forming
XX properties, or starch grain size or structure. This sequence represents
XX the soluble starch synthase isolated from wheat (Triticum aestivum L. cv.
XX Florida).
XX Sequence 756 AA;
Query Match 31.7%; Score 878; DB 21; Length 756;
Best Local Similarity 35.9%; Pred. No. 5.4e-67;
Matches 210; Conservative 78; Mismatches 164; Indels 74; Gaps 16;
QY 4 IVVYAAEVAPWSKGTGLGDTGGLPIELVKRGHRVMTIAPRY-----DQYADAWDTSVV 57
Db 141 IVFVTGEAAPYAKSGGLGDCVGLPIALAAARGHRVMMVMPRYLNGSSDKNYAKALYTAKH 200
QY 58 VDIM-----GEKRVFHSIKKGVHRWIDHPWFLAKVWGKTGSKLYGPRSGADYLDNHRKF 113
Db 201 IKIPCFGSGHEVTEFFEYRDNDVWVFDVDP-----SYHRPGS-LYGDNFGA-FGDNQFVY 253
QY 114 ALFCKAATEARVLPFGP---GECVCFVANDWSHVSALVPLKDEYQPKGQFTKAKSVLAI 170
Db 254 TLLCYACEAPLILLEGYIYQNCMFVNDWHASLVPIILAAKYRYPYGVYEDSKSLVI 313
QY 171 HNIAPQGRMWEAFKDTKLPOAFDKLAFSDGYAKVYATPMEDEKPLPTGKTYKKIN 230
Db 314 HNLAHQGVPEASTYDGLGLPPEWYGALEW-----VPEWARHRLDXG-----EAVN 360
QY 231 WLKGLIIADKLVISPNVAFETIADAAGGVELDTVI--RAKGIEGVNGMDIEWNEKT 288
Db 361 FLKGAVTADRTVTSQYSWEVTT-AEGGGLNELSSRSKSVLNGVINGIDINDWNETT 419
QY 289 DKFLSAPYDONSVAAGKAAKEALQAEGLPDPDPTAPLFAFICRLEEQGVVDIILAALPK 348
Db 420 DKCLPHHYSVDL--SGKAKKAEQLQELGLPVRDVPVPLIGFLGRDLYOKGIDLIKWAIFE 478

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 14:47:01 ; Search time 13.7251 Seconds
(without alignments)
3719.278 Million cell updates/sec

Title: US-09-980-771A-9
Perfect score: 2769
Sequence: 1 ALDVMVAEVAEVPWSKTGGL.....GGVATAKKEIKVPAEKIP 531

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1481.5	53.5	608	S43341	starch synthase (E
2	1457.5	52.6	607	YUPOX	starch synthase (E
3	1415.5	51.1	603	S61504	glycogen(starch) s
4	1412.5	51.0	608	T10906	starch synthase (E
5	1398.5	50.5	609	JQ0703	glycogen(starch) s
6	1397.5	50.5	609	S11481	glycogen(starch) s
7	1395.5	50.4	603	YUBHY	glycogen(starch) s
8	1390	50.2	610	F86453	granule-bound star
9	1362	49.2	615	YUTWY	glycogen(starch) s
10	1355.5	49.0	608	T14731	glycogen(starch) s
11	1353.5	48.9	605	S07314	glycogen(starch) s
12	902.5	32.6	641	T07668	glycogen(starch) s
13	873.5	31.5	626	JQ2322	starch synthase (E
14	872	31.5	732	T01208	starch synthase (E
15	871	31.5	752	S61505	starch synthase (E
16	866	31.3	610	T06280	glycogen(starch) s
17	861.5	31.1	622	T01414	probable starch sy
18	847.5	30.6	788	T07667	starch synthase (E
19	839	30.3	491	T07698	glycogen(starch) s
20	837	30.2	698	T01209	probable starch sy
21	684.5	24.7	477	B95130	glycogen synthase
22	683.5	24.7	477	T09800	starch synthase (E
23	677.5	24.5	484	S40051	starch synthase (E
24	669	24.2	483	C86712	starch synthase (E
25	652	23.5	480	H98228	glycogen synthase
26	652	23.5	480	H98228	glycogen synthase
27	651	23.5	486	A73057	glycogen synthase
28	637	23.0	477	A97176	glycogen synthase,
29	622	22.5	484	F82165	glycogen synthase

ALIGNMENTS

RESULT 1

S43341

starch synthase (EC 2.4.1.21) precursor - cassava

N:Alternate names: bacterial-glycogen synthase; glycogen synthase; starch synthase
C:Species: Manihot esculenta (cassava)

C>Date: 07-Sep-1994 #sequence_revision 03-Nov-1995 #text_change 16-Aug-2002

C:Accession: S43341

R:Salehuzzaman, S.N.I.M.; Jacobsen, E.; Visser, R.G.F.

Plant Mol. Biol. 23, 947-962, 1993

A:Title: Isolation and characterization of a cDNA encoding granule-bound starch synth

A:Reference number: S43341; MUID:94083565; PMID:8260633

A:Accession: S43341

A:Molecule type: mRNA

A:Residues: 1-608 <SAIL>

A:Cross-references: EMBL:X74160; NID:g437041; PIDN:CAA52273.1; PID:g437042

C:Genetics:

A:Gene: GBSS; waxy

A:Genome: nuclear

C:Function:

A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose produc

C:Superfamily: starch synthase

C:Keywords: amyloplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltra

F;1-78/Domain: transit peptide (amyloplast) #status predicted <NRP>

F;79-608/Product: ADPglucose-starch glucosyltransferase #status predicted <MAT>

Query Match 53.5%; Score 1481.5; DB 2; Length 608;

Best Local Similarity 55.0%; Pred. No. 6.4e-94;

Matches 299; Conservative 64; Mismatches 148; Indels 33; Gaps 9;

QY 2 LDIVMVAEVAEVPWSKTGGLGVDVTGGLPIELVKRHRVMTIAPRYDQYADAWDTSVVVDI- 60
 Db 82 MNLIFVGAEGVPSKGTGLGDLVGLPPAAARHRVMTVSPRYDQYADAWDTSVSVEIK 141
 QY 61 MG---EKVRFHSHKGVHVRWIDHPWFLAKVWGKSGSKLYGRPSGADYLDNHRKRALFC 117
 Db 142 IGRIETVRFHSHKGVHVRWIDHPWFLAKVWGKSGSKLYGRPSGADYLDNHRKRALFC 201
 QY 118 KAAIEAARVLPF-----GP-GEDCVFVANDHSHALVPVLLKDEYQPKQKFAKSVLA 169
 Db 202 LAALAPRVNLNSSKNFSGPYGEEVAFANDWHHTALLPCYLKAIYQPMGIYKHAKVAF 261
 QY 170 IHNIAFGRMWEAEAFKDTKLPAQAAFDKLAFCDSQYAKVYTEATPMEEDERPLLTGKYTKI 229
 Db 262 IHNIAFGRMWEAEAFKDTKLPAQAAFDKLAFCDSQYAKVYTEATPMEEDERPLLTGKYTKI 304
 QY 230 NMLKGGITAAADKLVTVSPNYATEIAADAAGGVELDTVIRAKGIEGVINGMDIEWNPKTD 289
 Db 305 NNMKAGILESRLVITVSPYAAQEVISGVERGVELDNFIRKTIAGIINGMDIQWNPVTD 364
 QY 290 KFLSAPYDONSYYACKAAKEALQALGLPVPDPTAFIPAFIGRLEKQKGVDTIILALPKI 349
 Db 365 KYVIDHYDATTVMADKAPLLKEALQAEVGLPVDNRNPLIGFGRLEKQKGVDTIILALPKI 423

starch (bacterial
probable starch sy
starch synthase (E
probable starch sy
starch synthase (E
glycogen synthase
glycogen synthase
starch synthase (E
starch synthase (E
soluble starch syn
hypothetical prote
glycogen synthase
glycogen synthase
glycogen (starch)
hypothetical prote

```

QY 350 LATPKVQIATILGTGKAAAYEKLVAIGTKYKRAKGVKFSAPLAHMLTAGADFMVPSRF 409
Db 424 LVEHNVQIVILGTGKKKEKQIEHLEVLVDPDKARGVAKENVPLAHMITAGADFMVPSRF 483
QY 410 EPCGLIQLHAMHYGTPVAVSTGGGLVDVTKEGVTGFHMGALNP--DKLDEADADALAIV 467
Db 484 EPCGLIQLHAMRYGTPVAVSTGGGLVDVTKEGVTGFHMGALNP--DKLDEADADALAIV 543
QY 468 RRASEVFAGGRPPEMVANCSODLSWSKPAQKWEGLLEEVYKGVATAKKEIKVPVA 527
Db 544 ARALGTATATAALREMILNCMAQDLSWKGPARMWEKMLLDLEV-TGSEPTGEGETIAPLAK 602
QY 528 EKIP 531
Db 603 ENVP 606

RESULT 2
YUPOY
starch synthase (EC 2.4.1.21) precursor - potato
N:Alternate names: starch synthase
C:Species: Solanum tuberosum (potato)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 02-Aug-2002
C:Accession: S16555; S24392; S26060; S26061
R:Van der Leij, F.R.; Visser, R.G.F.; Ponstein, A.S.; Jacobsen, E.; Feenstra, W.J.
Mol. Gen. Genet. 228, 240-248, 1991
A:Title: Sequence of the structural gene for granule-bound starch synthase of potato (Sc)
A:Reference number: S16555; MUID:91360072; PMID:1886609
A:Accession: S16555
A:Molecule type: protein
A:Residues: 1-607 <LEI>
A:Cross-references: EMBL:X58453; NID:g21470; PIDN:CAA41359.1; PID:g21471
A:Note: the authors translated the codon AAC for residue 453 as Gly and GCT for residue
A:Accession: S24392
A:Molecule type: protein
A:Residues: 78-92, 'X', '94-98, 'XXX', 102, 'XX', 105-107 <LE2>
R:Rohde, W.; Becker, D.; Kull, B.; Salamini, F.
J. Genet. Breed. 44, 311-315, 1990
A:Title: Structural and functional analysis of two waxy gene promoters from potato.
A:Reference number: S26060
A:Accession: S26060
A:Molecule type: DNA
A:Residues: 1-43 <ROH1>
A:Cross-references: EMBL:X52416; NID:g21613; PIDN:CAA36667.1; PID:g21614
A:Experimental source: cv. Granola, clone G1
A:Accession: S26061
A:Molecule type: DNA
A:Residues: 1-43 <ROH2>
A:Cross-references: EMBL:X52417; NID:g21615; PIDN:CAA36668.1; PID:g21616
A:Experimental source: cv. Granola, clone G28
C:Genetics:
A:Gene: waxy
A:Genome: nuclear
A:Introns: 111/3; 138/3; 171/3; 201/3; 223/1; 256/3; 293/2; 374/3; 433/3; 497/3; 526/3;
C:Function:
A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producing e
A:Pathway: starch biosynthesis
A:Superfamily: starch synthase
C:Keywords: amyloplast; glycogen/starch biosynthesis; glucosyltransferase; hexosyltransf
F:1-77/Domain: transit peptide (amyloplast) #status predicted <INP>
F:78-607/Product: ADPglucose-starch glucosyltransferase #status experimental <MAT>

Query Match 52.6%; Score 1457.5; DB 1; Length 607;
Best Local Similarity 54.0%; Pred. No. 2.8e-92;
Matches 295; Conservative 67; Mismatches 143; Indels 41; Gaps 10;

QY 2 LDIWVAEAVAPWSKTGGLGVTGCLPIELVKRGHVMVTIAPRYDQYADAMDTSVVVDI- 60
Db 81 MNLIFVGTVEVGPWSKTGGLGVTGCLPIELVKRGHVMVTIAPRYDQYADAMDTSVVVEVK 140
QY 61 MG---EKVRYFHSIKKGVHVRWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKRALFC 117

```

```

Db 141 VGSISIVRPFHCYKRCVDRVFDVHPMFLEKVVWGTGSKIYGPAGLDYLDNELRFSLLC 200
QY 118 KAAIEAARYLPF-----GP-GECCVFVANDKHSALVPLLKDETOPQKGQFTKAKSVLA 169
Db 201 QAALEAPKVLNLSNSSFSPGYGEDVLFITANDWHTALIPCYLKSMTQSRGIYLNKAVFC 260
QY 170 IHNIAFGQRMWEAFKDTIKPQAAFDKLAFCGAYKYTYTEATPMEDEKPPPLGKTKYKKI 229
Db 261 IHNIAFGQRFSTFDFLLNLPDEFGRGSDFDICYEK-----PVKG---RKI 303
QY 230 NMLKGLIADLVTVSPNYATEIAADAAGGVDELTVIRAKGLEGVINGMDIEWPKTD 289
Db 304 NMKAGILSHRVTVVSPYTAQELVSNADKGVSLDSVLRKTCITGIVNGMDTQWNPATD 363
QY 290 KFLSAPYDQNSVYAGKAAKEALQAEGLDVPDTPAFLEAFIGLEEQKGVDDIILALPKI 349
Db 364 KYTDVKYDITVMDAKPPLLKEALQAAVGLPVDKPKIPLIGFIRGLEEQKSDIILVAIHKF 423
QY 350 LATPKVQIATILGTGKAAAYEKLVAIGTKYKRAKGVKFSAPLAHMLTAGADFMVPSRF 409
Db 424 IGL-DWQIVVLGTGKKEFEQEIQLQEVLYPNKAGVAKFNVPLAHMITAGADFMVPSRF 482
QY 410 EPCGLIQLHAMHYGTPVAVSTGGGLVDVTKEGVTGFHMGALNP--DKLDEADADALAIV 467
Db 483 EPCGLIQLHAMRYGTPVAVSTGGGLVDVTKEGVTGFHMGAFNVBCDVPDADVLKIVTV 542
QY 468 RRASEVFAGGRPPEMVANCSODLSWSKPAQKWEGLLEEVYKGVATAKKEIKVPVA 519
Db 543 ARALVYGTALFAEMIKNCMEELSKPEAKKNTLILGLGASGSEPGVEGE-EIAPLAK 601
QY 520 EEIKVP 525
Db 602 ENVPAT 607

RESULT 3
S61504
glycogen(starch) synthase (EC 2.4.1.11) isoform I precursor - garden pea
N:Alternate names: glycogen(starch) synthase
C:Species: Pisum sativum (garden pea)
C:Date: 20-Jul-1996 #sequence_revision 07-Feb-1997 #text_change 16-Aug-2002
C:Accession: S61504; S72372
R:Dry, J.; Smith, A.; Edwards, A.; Bhattacharyya, M.; Dunn, P.; Martin, C.
Plant J. 2, 193-202, 1992
A:Title: Characterization of cDNAs encoding two isoforms of granule-bound starch synt
A:Reference number: S61504; MUID:93251108; PMID:1302049
A:Accession: S61504
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-603 <DRY>
A:Cross-references: EMBL:X88789; NID:g887570; PIDN:CAA61268.1; PID:g887571
A:Accession: S72372
A:Molecule type: protein
A:Residues: 76-77, 'X', 79, 81-88 <DRW>
C:Superfamily: starch synthase
C:Keywords: glucosyltransferase; hexosyltransferase
F:1-75/Domain: signal sequence #status predicted <SIG>
F:76-603/Product: glycogen(starch) synthase isoform I #status experimental <MAT>

Query Match 51.1%; Score 1415.5; DB 2; Length 603;
Best Local Similarity 51.8%; Pred. No. 2.2e-89;
Matches 284; Conservative 79; Mismatches 140; Indels 45; Gaps 12;

QY 2 LDIWVAEAVAPWSKTGGLGVTGCLPIELVKRGHVMVTIAPRYDQYADAMDTSVVVDI- 60
Db 77 MSLVFGVAGVGPWSKTGGLGVTGCLPIELVKRGHVMVTIAPRYDQYADAMDTSVVVEVK 136
QY 61 MGEK---VRYFHSIKKGVHVRWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKRALFC 117
Db 137 VGDKTETVRFHCYKRGVDRVFDVHPMFLEKVVWGTGSKIYGPAGLDYLDNELRFSLLC 196
QY 118 KAAIEAARYLPF-----GP-GECCVFVANDKHSALVPLLKDETOPQKGQFTKAKSVLA 169

```

Db 197 QAALEAPRVNLNSSKIFSGPYGEDVIFVANDWHSALIPCLYKSMVSKGLYKNAKVAFC 256
 QY 170 IHNIAFOGRWEEAFKPD---TKLPQAAEDKLFASDGYAKVYTEATPMEDEKPPPLTGKTY 226
 Db 257 IHNIAOGR---NAFSDFSLNLPDEFSSFDIDYK-----PCBG--- 296
 QY 227 KKNLWKGITIAADKLVTVSPNYATEIAADAAGGVDELIVIRAKGIEGIVNGMDTEWNP 286
 Db 297 KKNWKGILSDQVFTVSPNYATEIAADAAGGVDELIVIRAKGIEGIVNGMDTEWNP 286
 QY 287 KIDKLSAPYDQNSVYAGKAAKEALQAEGLPVDPTAPLFAFIRLEBQKGVDIILALPKI 349
 Db 357 QTDRIYDHYNETTTEAKPLKGLQAEIGLPVDSIPPLIGFIRLEBQKGVDIILALPKI 349
 QY 347 PKILATPKVQIALGTGAAYEKLVNAIGTKYKGRAGKGVKFSAPLAHMLTAGADFMVLP 406
 Db 417 AK-FADENVQIVLGTGKIMKQKQEVLEEKYPGKAIGITKFNPSLAHMLTAGADFMVLP 406
 QY 407 SRFEPCGLIOLHAMHYGTVPVVASTGGLVDTVKEGTVGHMGAALNP---DKLDEADALALA 464
 Db 476 SRFEPCGLIOLHAMHYGTVPVVASTGGLVDTVKEGTVGHMGAALNP---DKLDEADALALA 464
 QY 465 ATVRASEVFAAGRPYEMVANCISODLSWSKPAQKWEGLLEEV-----VYKGG--VATA 517
 Db 536 ATVRKALTYGTQAKQIILNCMAQNFSSWKKPAKLEKALLNLEVTGNVAGIDGDEIAPL 595
 QY 518 KKEIKVP 525
 Db 596 AKENWATP 603

RESULT 4

T10906
 A: glycosyltransferase (EC 2.4.1.21) - sweet potato
 N: Alternate names: starch synthase
 C: Species: Ipomoea batatas (sweet potato)
 C: Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Aug-2002
 C: Accession: T10906
 R: Wang, S.J.; Yeh, K.W.; Tsai, C.Y.
 A: Reference number: 217212
 A: Accession: T10906
 A: Status: preliminary; translated from GB/EMBL/DBJ
 A: Molecule type: mRNA
 A: Residues: 1-608 <WAN>
 A: Cross-references: EMBL:U44126; NID:gl172158; PID:gl172159
 A: Experimental source: cv. Tainong, tuberous root
 C: Genetics:
 A: Gene: SS67
 C: Function:
 A: Description: catalyzes the alpha-1,4-glucosylation of starch by ADP-glucose producing
 A: Pathway: starch biosynthesis
 C: Superfamily: starch synthase
 C: Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 51.0%; Score 1412.5; DB 2; Length 608;
 Best Local Similarity 53.5%; Pred. No. 3.5e-89;
 Matches 292; Conservative 61; Mismatches 152; Indels 41; Gaps 10;

QY 2 LDIVMVAEAPWSPKSGGLDVTGGGLPIELVKRHRVMTIAPRYDQYADAWDTSVVVDI- 60
 Db 82 MNLVFGCEGPMCKTGGGLDVLGGLPPALAAARGHRVMTVCPRYDQYKDAWETCVVVEPQ 141
 QY 61 MG---EKVRVFSIKKGVHVRWIDHPWFLAKVWGKSGKLYGPRSGADYLDNHRKRALFC 117
 Db 142 VGRDIEVRFHFKYKGVDRVDFVDFHMFLEKVGWGTGSMLYGPKAGKDYKDNQRLRSLIC 201
 QY 118 KAAIEAARVLPF-----GP-GEDCVFVANDWHSALVPVLLKDEYKQKGFKAQSVLA 169
 Db 202 QAALEAPRVNLNSSKIFSGPYGEDVIFVANDWHSALVPVLLKDEYKQKGFKAQSVLA 169
 QY 170 IHNIAFOGRWEEAFKPDTKLPQAAFDKLFASDGYAKVYTEATPMEDEKPPPLTGKTYK 229

Db 262 IHNIAOGRFAFSDFSLNLPDEYKGSFDFIDYK-----PVKG---RKI 304
 QY 230 NMLKGGIILAADKLVTVSPNYATEIAADAAGGVDELIVIRAKGIEGIVNGMDTEWNP 289
 Db 305 NMLKGGIILAADKLVTVSPNYATEIAADAAGGVDELIVIRAKGIEGIVNGMDTEWNP 289
 QY 290 KFLSAPYDQNSVYAGKAAKEALQAEGLPVDPTAPLFAFIRLEBQKGVDIILALPKI 349
 Db 365 KYLAVKYDITVMOAKPPLKKEALQAAVGLPVDNIFLIGFIRLEBQKGVDIILALPKI 349
 QY 350 LATPKVQIALGTGAAYEKLVNAIGTKYKGRAGKGVKFSAPLAHMLTAGADFMVLP 409
 Db 425 ISM-DVQIILGTGKKEKFEQLEFVMPDPKARGVAKENVPLAHMHTAGADFMVLP 483
 QY 410 EPCGLIOLHAMHYGTVPVVASTGGLVDTVKEGTVGHMGAALNP---DKLDEADALALA 467
 Db 484 EPCGLIOLHAMHYGTVPVVASTGGLVDTVKEGTVGHMGAALNP---DKLDEADALALA 467
 QY 468 RRASEVFAAGRPYEMVANCISODLSWSKPAQKWEGLL-----FEVYKGGVATAKK 519
 Db 544 GRALATYGTLAFTTEMIKNCMSQELSKWGPKNWEIVLLSIGVAGSEPGVEGE-ETAPLAK 602
 QY 520 BEIKVP 525
 Db 603 ENWATP 608

RESULT 5

JQ0703
 A: glycogen(starch) synthase (EC 2.4.1.11) - rice
 N: Alternate names: granule-bound starch synthase
 C: Species: Oryza sativa (rice)
 C: Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Aug-2002
 C: Accession: JQ0703
 R: Wang, Z.Y.; Wu, Z.L.; Xing, Y.X.; Zheng, F.G.; Guo, X.L.; Zhang, W.G.; Hong, M.M.
 A: Nucleic Acids Res. 18, 5898, 1990
 A: Title: Nucleotide sequence of rice waxy gene.
 A: Reference number: JQ0703; MUID:91016948; PMID:2216792
 A: Accession: JQ0703
 A: Status: translation not shown
 A: Molecule type: DNA
 A: Residues: 1-609 <WAN>
 A: Cross-references: EMBL:X53694
 A: Experimental source: strain subsp. Japonica Hengfeng
 C: Genetics:
 A: Gene: waxy
 A: Introns: 114/1; 174/1; 204/1; 225/2; 259/1; 295/3; 377/1; 436/1; 500/1; 529/1
 C: Function:
 A: Description: catalyzes the alpha-1,4-glucosylation of starch by UDP-glucose produc
 C: Superfamily: starch synthase
 C: Keywords: chloroplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltr

Query Match 50.5%; Score 1398.5; DB 2; Length 609;
 Best Local Similarity 51.2%; Pred. No. 3.2e-88;
 Matches 280; Conservative 77; Mismatches 147; Indels 43; Gaps 9;

QY 2 LDIVMVAEAPWSPKSGGLDVTGGGLPIELVKRHRVMTIAPRYDQYADAWDTSVVVDI- 60
 Db 83 MNLVFGCEGPMCKTGGGLDVLGGLPPALAAARGHRVMTVCPRYDQYKDAWETCVVVEPQ 142
 QY 61 MG---EKVRVFSIKKGVHVRWIDHPWFLAKVWGKSGKLYGPRSGADYLDNHRKRALFC 117
 Db 143 VGRDIEVRFHFKYKGVDRVDFVDFHMFLEKVGWGTGSMLYGPKAGKDYKDNQRLRSLIC 202
 QY 118 KAAIEAARVLPF-----GP-GEDCVFVANDWHSALVPVLLKDEYKQKGFKAQSVLA 169
 Db 203 QAALEAPRVNLNSSKIFSGPYGEDVIFVANDWHSALVPVLLKDEYKQKGFKAQSVLA 169
 QY 170 IHNIAFOGRWEEAFKPDTKLPQAAFDKLFASDGYAKVYTEATPMEDEKPPPLTGKTYK 229
 Db 263 IHNISYOGRAFEYDPELNLSEFRSFDIDG-----DTPVEG-----RKI 305
 QY 230 NMLKGGIILAADKLVTVSPNYATEIAADAAGGVDELIVIRAKGIEGIVNGMDTEWNP 289

Db 306 NWKAGILEADRVLTSPYAEELISSIARGCELDNIMRLTITGITVINGMDVSEWDPKSD 365
 QY 290 KFLSAPYDQNSVYAGKAAKEALQAEGLPVDPTAFIPAGHLEBOKGVDIILALPKI 349
 Db 366 KYITAKYDATTAEAKALKEALQAEAGLPVDRKIPLIAPFAGLEBOKGPVMAAAIPEL 425
 QY 350 LATPKVQIAILGTGKAYEKLVAIGTKYGRAGKGVVKSAPLAHMLTAGADEMLVPSRF 409
 Db 426 M-QEDVQIIVLGTGKAYEKLVAIGTKYGRAGKGVVKSAPLAHMLTAGADEMLVPSRF 484
 QY 410 EPCGLIOLHAMHYGTPVVAASGTGLVDTVKEGTVGFHMGALNPD--KLDEADADALAATV 467
 Db 485 EPCGLIOLQGMRYGTACACASTGGLVDTVIEGKTFGHMGRSLVSDGKVVESDVKKVAATL 544
 QY 468 RRASEVAGGRYPPEMVAANCISQDLSWSKPAQKNGLEEVYVYKGG-----VATAK 518
 Db 545 KRAIKVGTTPAYEEMVRCNMQDLSWKGPAKNWENVL--LGLGVAGSAPGIESDEIAPLA 602
 QY 519 KEIKVVP 525
 Db 603 KENVAAP 609

RESULT 6

S11481
 glycogen(starch) synthase (EC 2.4.1.11) precursor - rice
 N:Alternate names: starch synthase; waxy protein
 C:Species: Oryza sativa (rice)
 C:Date: 05-Mar-1995 #sequence_revision 10-Nov-1995 #text_change 16-Aug-2002
 C:Accession: S11481; S22519; S30485; PC2190; JQ2224
 R:Wang, Z.Y.; Wu, Z.L.; Xing, Y.Y.; Zheng, F.G.; Guo, X.L.; Zhang, W.G.; Hong, M.M.
 Nucleic Acids Res. 18, 5898, 1990
 A:Title: Nucleotide sequence of rice waxy gene.
 A:Reference number: JQ0703; MUID:91016948; PMID:2216792
 A:Accession: S11481
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-609 <WAN>
 A:Cross-references: EMBL:X53694; NID:g577598; PIDN:CAA37732.1; PID:g577599
 R:Okagaki, R.J.
 Plant Mol. Biol. 19, 513-516, 1992
 A:Title: Nucleotide sequence of a long cDNA from the rice waxy gene.
 A:Reference number: S22519; MUID:92322986; PMID:1377969
 A:Accession: S22519
 A:Molecule type: mRNA
 A:Residues: 1-609 <OKA>
 A:Cross-references: EMBL:X62134; NID:g20402; PIDN:CAA44065.1; PID:g20403
 R:Hirano, H.Y.; Sano, Y.
 Plant Cell Physiol. 32, 989-997, 1991
 A:Title: Molecular characterization of the waxy locus of rice (Oryza sativa).
 A:Reference number: S30485
 A:Accession: S30485
 A:Status: preliminary; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-609 <HIR>
 A:Cross-references: EMBL:X58228; NID:g20400; PIDN:CAA41186.1; PID:g20401
 A:Accession: PC2190
 A:Molecule type: protein
 A:Residues: 78-113 <HIZ>
 A:Experimental source: leaf, cDNA POSLHC2120
 C:Comment: This protein is involved in amylose synthesis in the rice endosperm.
 C:Genetics:
 A:Gene: waxy
 A:Introns: 113/3; 140/3; 173/3; 203/3; 225/1; 258/3; 295/2; 376/3; 435/3; 499/3; 528/3;
 C:Function:
 A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producing
 A:Pathway: starch biosynthesis
 C:Superfamily: starch synthase
 C:Keywords: amyloplast; chloroplast; glycogen/starch biosynthesis; glycosyltransferase;
 F:1-77/Domain: transit peptide (amyloplast) #status predicted <TNP>
 F:78-609/Product: glycogen (starch) synthase #status experimental <MAT>

Query Match 50.5%; Score 1397.5; DB 2; Length 609;
 Best Local Similarity 51.4%; Pred. No. 3.3e-88;
 Matches 281; Conservative 76; Mismatches 147; Indels 43; Gaps 9;
 QY 2 LDIVWAAEAVAPSKTGGGLDVTGGLPIELVKRGHRVMTIAPRYDQYADAMDTSVVVDI- 60
 Db 83 MNVVFVGAEMAPNSKTGGGLDVTGGLPPMAANGHRVMTISPRDQYKDAWDTSVVAEIK 142
 QY 61 ---MGEKRVYFHSIKKGVRHVWIDHPWFLAKVWGTGSKLYGPRSGADYLDNHRKFALEFC 117
 Db 143 VADRYENRVFHCYKRGVORVFDHPSELEKVGWTKGEKIYGPDTGVDTYKDNQMRFSLLC 202
 QY 118 KAAIEAARVL-----PFGP---GEDCVFVANDHISALVPVLLKDEYQKQGTAKKSVLA 169
 Db 203 QAALEAPRIILNPNPYFKGTYGEDVVFVFCNDWHTGPLASLYLKNYQPNGYIRNAKVAFC 262
 QY 170 IHNIAFOGRWEEAFKDTKLQAAFDKLAESDGYAKVYVTEATPMEDEKPLTGTGYKKI 229
 Db 263 IHNISYQGRFAFEDYPELNLSEFSSFDIDGY-----DTPVEG-----RKI 305
 QY 230 NWLKGIIADKLVTSVPNYATEIAADAAGGVVELDTVIRAKGIEGIVNGMDIEWNPKTD 289
 Db 306 NNMKAGILEADRVLTSPYAEELISGIARGCELDNIMRLTITGITVINGMDVSEWDPKSD 365
 QY 290 KFLSAPYDQNSVYAGKAAKEALQAEGLPVDPTAFIPAGHLEBOKGVDIILALPKI 349
 Db 366 KYITAKYDATTAEAKALKEALQAEAGLPVDRKIPLIAPFAGLEBOKGPVMAAAIPEL 425
 QY 350 LATPKVQIAILGTGKAYEKLVAIGTKYGRAGKGVVKSAPLAHMLTAGADEMLVPSRF 409
 Db 426 M-QEDVQIIVLGTGKAYEKLVAIGTKYGRAGKGVVKSAPLAHMLTAGADEMLVPSRF 484
 QY 410 EPCGLIOLHAMHYGTPVVAASGTGLVDTVKEGTVGFHMGALNPD--KLDEADADALAATV 467
 Db 485 EPCGLIOLQGMRYGTACACASTGGLVDTVIEGKTFGHMGRSLVSDGKVVESDVKKVAATL 544
 QY 468 RRASEVAGGRYPPEMVAANCISQDLSWSKPAQKNGLEEVYVYKGG-----VATAK 518
 Db 545 KRAIKVGTTPAYEEMVRCNMQDLSWKGPAKNWENVL--LGLGVAGSAPGIESDEIAPLA 602
 QY 519 KEIKVVP 525
 Db 603 KENVAAP 609

RESULT 7

YUBHY

glycogen(starch) synthase (EC 2.4.1.11) precursor - barley
 N:Alternate names: starch synthase
 C:Species: Hordeum vulgare (barley)
 C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 02-Aug-2002
 C:Accession: S01727; S01728
 R:Rohde, W.; Becker, D.; Salamini, F.
 Nucleic Acids Res. 16, 7185-7186, 1988
 A:Title: Structural analysis of the waxy locus from Hordeum vulgare.
 A:Reference number: S01727; MUID:88303345; PMID:2970062
 A:Accession: S01727
 A:Molecule type: DNA
 A:Residues: 1-603 <ROH>
 A:Cross-references: EMBL:X07931; NID:g19126; PIDN:CAA30755.1; PID:g295809
 A:Accession: S01728
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-603 <ROH2>
 A:Cross-references: EMBL:X07932; NID:g19128; PIDN:CAA30756.1; PID:g19129
 C:Genetics:
 A:Gene: waxy
 A:Introns: 106/3; 133/3; 166/3; 218/1; 251/3; 369/3; 429/3; 493/3; 522/3; 565/3
 C:Function:
 A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producin
 A:Pathway: starch biosynthesis
 C:Superfamily: starch synthase
 C:Keywords: amyloplast; starch biosynthesis; glycosyltransferase; hexosyltra